PREFACE.

The following thesis, submitted to the University of Tasmania for the degree of Doctor of Science, consists of two main sections, complete details of which may be found on pp. (ii)-(v). The parts marked [1] - [5] represent published papers, whilst [6] - [8] have been submitted for publication to the Acta Mathematica and the Annals of Mathematics respectively. The rest consists of a number of hitherto unpublished papers.

The work was carried out independently by the candidate, and the various ideas underlying the construction of the algebraic optics here presented, as well as the methods of dealing with invariant action principles are due to him. Further details may be found on pp. 3, 34 ff., 149 ff., etc.

Of two papers photostatic copies of proofs are given, as reprints have not yet been received. In accordance with the rules governing the degree, two papers on the Axiomatic Treatment of the Second Law of Thermodynamics, (due to appear in the American Journal of Physics in December 1948), have also been included.

H Buchdahl

September, 1948.

Note: Throughout this thesis, underlined symbols are to be read as follows:

Red	......	Clarendon type,
Green	......	German (Gothic) type,
Black	......	Italics.

CONTENTS.

SECTION A : OPTICS.

ALGEBRAIC METHODS FOR THE DETERMINATION OF THE GEOMETRICAL HIGHER ORDER ABERRATIONS OF OPTICAL SYSTEMS.

contd.

- - - - - - - -

It has been published as: Buchdahl, H. A., 1946. The algebraic analysis of the higher-order aberrations of optical systems : tangential aberrations of a system of coaxial spherical refracting surfaces, Proceedings of the Physical Society, 58(5), 545-575

THE ALGEBRAIC THEORY AND CALCULATION OF THE GEOMETRICAL HIGHER ORDER ABERRATIONS OF OPTICAL SYSTEMS.

I: Aberrations of systems of coaxial spherical refracting surfaces.

by H. A. Buchdahl.

Dept. of Physics, University of Tasmania.

Abstract: Defining the aberration of a ray in terms of the components of the displacement of its point of intersection with a chosen plane of reference from an ideal image point in the same plane, a theory of higher order aberrations is developed by means of elementary algebraic methods. The treatment rests in principle upon the study of the departure from exact linearity of the relations connecting the variables specifying the ray at the different surfaces of an optical system. This leads in practice to the possibility of computing the exact aberration coefficients to any order desired without the necessity of any trigonometrical tracing. Unrestricted changes in the position of object, diaphragm or plane of reference are easily dealt with in practice, the aberration coefficients being pure constants of the system. A number of special problems are briefly considered, such as the addition of different optical systems, and the aberrations of reversed systems.

As a practical example the computations necessary for the determination of the complete exact primary and secondary

aberrations, and of the tertiary spherical aberration are given in detail for the case of a Cooke Triplet illustrating the relatively small amount of labour involved. A number of diagrams illustrate the good agreement between the aberrations predicted by the present method and those obtained by means of strict trigonometrical tracing.

PART I. THEORY IN GENERAL.

§ 1. General Introduction.

(a) In the present paper, which is a continuation of the author's previous work (Buchdahl, 1946), a general algebraic theory of the monochromatic aberrations of systems of coaxial spherical refracting surfaces will be developed. The previous restriction to tangential rays is now removed, all rays capable of passing through the system being allowed. Since a ray is now specified by four co-ordinates instead of only two the theory may be expected to be somewhat more complex. Accordingly certain changes have been made in the method and notation used previously for the sake of simplicity of development as well as of uniformity of notation. In particular, considerable use has been made of an extension of the method outlined in §14 of the paper referred to above. This leads in practice to the possibility of computing the exact aberration coefficients, suitably defined, by means of an iterative method to any order desired. The calculation of the full quaternary or higher aberrations will usually be too lengthy to be worthwhile. The computing scheme gives the individual contributions of the different surfaces, and may be arranged so as to allow of a simple but complete investigation of unrestricted changes in the positions of object, diaphragm and plane of reference. It may be remarked that Herzberger states (Herzberger, 1931 (b)) that equations for the dependence of aberrations on object position are given by him for the first time (loc. cit. p. 115); but he confines himself entirely to a theoretical discussion of the primary terms. It should be stressed that aberrations of order higher than the first depend somewhat on the co-ordinates used unless those of lower order vanish identically: the differences between the results obtained with different co-ordinates are implicitly compensated by the appropriate remainder terms.

(b) In principle the simplicity of the method appears to arise mainly from three factors: first, the use of quasi-invariants, i.e. expressions which reduce to optical invariants in the paraxial limit (§3). These automatically render the derivation of explicit summation theorems unnecessary. Secondly, the choice of special systems of co-ordinates, below called canonical, which consist of two particular pairs of so-called linear co-ordinates (§4). Thirdly, the construction of expressions which closely resemble the linear relations of paraxial theory and which clearly indicate the departures from linearity; and it is these which make iteration possible (§§4b. and 5).

When canonical co-ordinates are used the aberration coefficients are pure constants of the system, in the sense that they do not involve the positions of the object, diaphragm or plane of reference. However, if desired, general linear co-ordinates may be used (§12.), such as pairs of co-ordinates in object and image space, etc. Special attention is paid below to co-ordinates closely resembling canonical co-ordinates (§4c.), the use of which makes it possible directly to employ pairs of direction cosines in the object space and image space respectively in order to specify the ray. This may be of some importance in any problems involving the use of the angle eikonal. When the system is telescopic this particular set of co-ordinates is forbidden. But we may be certain that canonical co-ordinates are always allowed.

(c) It may be remarked that the length of a synthetic algebraic theory as compared with a development along the lines of Hamilton's analytic method (e.g. Steward, 1928 (a)) is somewhat deceptive. This is mainly due to the fact that the aims of these rival theories do not coincide. No doubt Hamilton's method is best suited to the most general abstract development of geometrical optics. In the present paper, however, an attempt has been made to find the best compromise between brevity and elegance on the one hand and practical usefulness on the other: in particular as regards the actual calculation of the aberrations of higher order in the presence of those of lower order; whilst some writers on the subject seem to pay little attention to this problem (e.g. Herzberger, 1931 (a)). Our use of co-ordinates lying entirely in the object space contributes greatly to the ease with which practical problems may be handled (v. Herzberger, 1931 (c)); and no constants are introduced which cannot be straightforwardly determined in practice by the methods described below. Restriction to primary aberrations allows certain modifications to be made which render the method of quasi-invariants very valuable from a didactic point of view. In particular, the usual expressions for the Seidel aberrations (§16) may be derived in a manner probably more suitable for undergraduate teaching than the usual developments (e.g. Whittaker, 1915(a), or Conrady, 1929(a))

It will be noticed that virtually all the lengthy and cumbrous parts of the theory arise in the attempt to make the necessary calculations sufficiently brief to be useful in practice. The simplicity of the equations (20.41) and (20.42) in conjunction with (21.2) for instance shows that the somewhat laborious developments in Part III are amply justified.

§ 2. Notation.

(a) For convenient reference a table containing most of the symbols employed in the text is given at the end of this paper in §27.

The quantities $L(\equiv L_{sk})$, $H(\equiv -H_{sk})$, U_y, U_z, U, I, U_c, O, N, r, d' are essentially the same as those used by Conrady (Conrady, 1929 (b)); although these are not used finally they are employed at first in order to make the transition to canonical variables more transparent to the practical worker. The equations determining the passage of a skew ray through a surface are

$$\tan U_c = \frac{H}{L-r} \tag{2.1}$$

$$\tan O = \frac{\tan U_z}{\sin(U_y - U_c)} \tag{2.2}$$

$$\sin U = \frac{\sin U_z}{\sin O} \tag{2.3}$$

$$\sin I = \frac{H \sin U}{r \sin U_c} = \frac{(L-r)\sin U}{r \cos U_c} \tag{2.4}$$

$$\Delta(N \sin I) = 0 \tag{2.5}$$

$$\Delta(I + U) = 0 \tag{2.6}$$

together with the equations (2.1) - (2.4), with H, L, U_y, U_z, U, I replaced by the corresponding dashed symbols. The transfer equations are

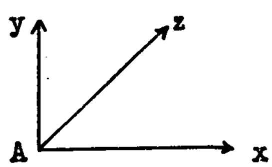
$$(U_y)_{j+1} = (U'_y)_j \tag{2.71}$$

$$(U_z)_{j+1} = (U'_z)_j \tag{2.72}$$

$$H_{j+1} = H'_j \tag{2.73}$$

$$L_{j+1} = L'_j - d'_j \tag{2.74}$$

Notice that we have taken H to be positive when it lies below the x - axis, the co-ordinate system being left-handed, viz.,



The x axis lies in the direction of the axis of symmetry of the system. A coincides with the pole of the refracting surface in question, the x, y plane coinciding with the tangential plane.

As regards the notation employed it may be remarked that the number of symbols used has been reduced to a minimum by the frequent and, above all, systematic use of affixes. Even so, duplication of certain symbols was unavoidable in some cases. But it is hoped that this will not lead to confusion.

(b) The numbered equations in the author's previous paper (Buchdahl, 1946) will be distinguished by the letter T; thus (T 5.8). A homogeneous polynomial of degree $2n+1$ in a set of linear variables (v. §4) will be said to be of the n th order. The phrase "correct to the n th order" as applied to any expression means that it holds when terms of order exceeding n are neglected. By analogy with usual mathematical notation the symbol $O(n)$ is used to represent a term or terms the degree of which is not less than n.

§3. Definition of the Aberrations of a Ray. Order. Quasi-Invariants.

(a) Let us take all paraxial rays always to lie in the tangential plane. Given two arbitrary paraxial rays (y_1, u_1) and $(\bar{y}_1, \bar{u}_1)$ the expression

$$\lambda = N(\ell - \bar{\ell})u\bar{u} \tag{3.1}$$

is an optical invariant according to (T 4.6). That is, for all j

$$\Delta\lambda_j = \lambda'_j - \lambda_j = 0 \tag{3.21}$$

and
$$\lambda_{j+1} = \lambda'_j, \tag{3.22}$$

so that λ has the same value everywhere in the optical system.

Let O be an object point lying in the tangential plane, whilst the "axial point of the object O_o" is the perpendicular projection of O on to the principal axis. Then the position of O may be specified by the position ℓ_{o1} of O_o with respect to the first surface and by the distance $O_0O = \underline{h}_1$. Moreover it is convenient to distinguish quantities referring to an "axial ray" (a ray through O_0) by the additional subscript $\underline{o}$.

If the two rays in (3.1) be taken to be respectively an axial ray and a ray through the object point, then

$$\lambda_j = N_j(\ell_{oj} - \ell_j)u_{oj}\, u_j = N_j\, h_j\, u_{oj} \tag{3.31}$$

is an optical invariant (Lagrange, or Smith-Helmholtz invariant).

If the suffix $\underline{k}$ refers to the last surface throughout, the paraxial magnification of the system m' ($=1/\underline{m}$) is defined as

$$m' = \frac{h_k'}{h_1} = \frac{N_1 u_{o1}}{N_k' u_{ok}'}, \quad \text{by (3.31)} \tag{3.32}$$

$$= \frac{\mu}{\mu'}, \quad \text{say.} \tag{3.21}$$

Since the right-hand side of (3.32) depends only upon $\underline{\ell}_{o1}$ it follows that all rays through O on emerging from the system pass through a certain point O', usually referred to as the <u>ideal image point</u>, the position of which is defined by $\underline{\ell}_{ok}'$ and $\underline{h}_k'$. The plane perpendicular to the principal axis which passes through the image point O_0' conjugate to O_0 is the <u>paraxial</u> or <u>ideal image plane</u>; and O' is situated in the latter. Since $\underline{h}_k'$ is proportional to $\underline{h}_1$ a linear object lying in the object plane will have a perfect image lying in the ideal image plane.

(b) Considering now actual rays passing through the system these will not in general intersect the ideal image plane in the point O'. Let the ray intersect the jth "object plane" in the point given by the rectangular co-ordinates H_{yj}, H_{zj}. If the ray passed through the system according to the laws of paraxial imagery the co-ordinates of this point would be $\underline{h}_{yj}$, $\underline{h}_{zj}$. Then the components of the aberration of the ray before the jth surface are

$$\left.\begin{aligned} \varepsilon_{yj} &= H_{yj} - h_{yj} \\ \varepsilon_{zj} &= H_{zj} - h_{zj} \end{aligned}\right\} \quad (3.4)$$

Since the optical system is symmetrical we may without loss of generality take the object point to lie in the tangential plane. In that case $\underline{h}_{zj} = 0$, and $\underline{h}_{yj}$ will be written $\underline{h}_j$. The sign convention for H_y, and H_z is the same as that which applies to H. ($\underline{v}$. §2a).

In order to avoid unnecessary duplication of many equations we shall make use of the convention that a symbol in Clarendon type shall represent the two corresponding quantities referring to the $\underline{y}$ and $\underline{z}$ axes; thus $\underline{\varepsilon}_j$ stands for both ε_{yj} and ε_{zj}. Also the general suffix $\underline{j}$ will be omitted whenever confusion is not likely to arise thereby. Eq. (3.4) then reads simply

$$\underline{\varepsilon} = \underline{H} - \underline{h} \quad (3.41)$$

Corresponding to the quantity λ in (3.31) we may write down the two expressions

$$\underline{\Lambda} = N\, u_o\, \underline{H} \tag{3.5}$$

Then
$$\Delta\underline{\Lambda} = \Delta N\, u_o\, \underline{H} = \Delta N\, u_o \underline{\varepsilon} \tag{3.51}$$

$$\text{since} \quad \Delta N\, u_o\, \underline{h} = 0 \tag{3.511}$$

Between the two "components" of $\Delta N\, u_o \underline{\varepsilon}$ there exists the remarkable identity

$$\Delta N\, u_o \varepsilon_y = \sec U_c \operatorname{cotan} O \Delta N u_o \varepsilon_z \tag{3.6}$$

This may be proved as follows: from elementary considerations we have

$$\left.\begin{aligned} H_y &= (\ell_o - L) \tan U_y + H \\ H_z &= (\ell_o - L) \tan U_z \sec U_y \end{aligned}\right\} \tag{3.61}$$

Hence

$$\Delta N u_o \varepsilon_y - \sec U_c \operatorname{cotan} O \Delta N u_o \varepsilon_z$$

$$= \Delta N u_o \left\{ H + (\ell_o - L) \left[\tan U_y - \tan U_z \sec U_y \sec U_c \operatorname{cotan} O \right] \right\}$$

$$= \Delta N u_o \left\{ H + (\ell_o - L) \left[\tan U_y - \sin(U_y - U_c) \sec U_y \sec U_c \right] \right\}, \text{ by } (2.2)$$

$$= \Delta N u_o \left\{ (L - r) \tan U_c + (\ell_o - L) \tan U_c \right\}, \text{ by } (2.4)$$

$$= \tan U_c\, \Delta\, c_o = 0, \quad \text{by } (T\ 1.6),$$

Q.E.D.

(c) Now since $\underline{\Lambda}_{j+1} = \underline{\Lambda}'_j$ $\qquad (j = 1, 2, \ldots, k-1)$ (3.7)

we can write at once

$$\left.\begin{aligned} \mu' \underline{\varepsilon}_k' &= \underline{\Lambda}'_k = \underline{\Lambda}_1 + \sum_{j=1}^{k} \Delta\underline{\Lambda}_j \\ &= \underline{\Lambda}_1 + \underline{D}_k', \text{ say} \end{aligned}\right\} \quad (3.71)$$

which shows how $\underline{\varepsilon}_k'$ is made up of the <u>contributions</u> $\Delta\underline{\Lambda}_j$ of the individual surfaces. In order to fix our ideas we must at this stage agree on some definite method of specifying a particular ray. Accordingly at any surface let Y, Z be the co-ordinates of the point of intersection of the ray with the tangent plane to the pole of that surface; also let $(\alpha, -\beta, -\gamma)$ be the direction cosines of the ray, so that

$$\left.\begin{aligned} \alpha &= \cos U_y \cos U_z \\ \beta &= \sin U_y \cos U_z \\ \gamma &= \sin U_z \end{aligned}\right\} \quad (3.8)$$

Then we may specify the ray by the four co-ordinates Y_1, Z_1, V_1, W_1, where

$$\left.\begin{aligned} V &= \beta/\alpha \\ W &= \gamma/\alpha \end{aligned}\right\} \quad (3.81)$$

$\underline{\varepsilon}_k'$ may now be looked upon as a definite function of those four co-ordinates. To obtain this function in a closed form is out of the question in all but the simplest cases. Accordingly let $\Delta\underline{\Lambda}_j$ (and so $\underline{\varepsilon}_k'$) be expanded in a Taylor series. In virtue of the existence of an ideal image point the linear terms must be absent in this expansion. And since the system is symmetrical, $\Delta\underline{\Lambda}_j$ must transform in the same way as $\underline{Y}_1$ with respect to a rotation of the system. It follows that at any surface $\Delta\underline{\Lambda}$ will be of the form

$$\begin{aligned}\Delta\underline{\Lambda} = {} & a\underline{Y}_1\xi + \bar{a}\underline{V}_1\xi + b\underline{Y}_1\eta + \bar{b}\underline{V}_1\eta + c\underline{Y}_1\zeta + \bar{c}\underline{V}_1\zeta + s_1\underline{Y}_1\xi^2 \\ & + \bar{s}_1\underline{V}_1\xi^2 + s_2\underline{Y}_1\xi\eta + \bar{s}_2\underline{V}_1\xi\eta + s_3\underline{Y}_1\xi\zeta + \bar{s}_3\underline{V}_1\xi\zeta \\ & + s_4\underline{Y}_1\eta^2 + \bar{s}_4\underline{V}_1\eta^2 + s_5\underline{Y}_1\eta\zeta + \bar{s}_5\underline{V}_1\eta\zeta + s_6\underline{Y}_1\zeta^2 \\ & + \bar{s}_6\underline{V}_1\zeta^2 + t_1\underline{Y}_1\xi^3 + \bar{t}_1\underline{V}_1\xi^3 + t_2\underline{Y}_1\xi^2\eta + \dots\end{aligned} \tag{3.9}$$

where

$$\left.\begin{aligned}\xi &= Y_1^2 + Z_1^2 \\ \eta &= Y_1V_1 + Z_1W_1 \\ \zeta &= V_1^2 + W_1^2\end{aligned}\right\} \tag{3.91}$$

and the a, ā, ... are constants.

Clearly (3.9) consists of a series of homogeneous polynomials of degree 3, 5, 7, ... Accordingly we shall call the polynomial of degree $2n+1$, $(n = 1, 2, \dots)$ the nth order contribution (by the jth surface) to the aberration. The aberration of order n is then simply the sum of the contributions divided by μ'.

Thus, if we write

$$\left.\begin{aligned}A_j &= \sum_{\nu=1}^{j-1} a_\nu \\ A^j &= \sum_{\nu=1}^{j} a_\nu\end{aligned}\right\} \tag{3.92}$$

and similarly for all other coefficients, then

$$\underline{\varepsilon}_k' = \frac{1}{\mu'}\left\{A_k'\underline{Y}_1\xi + \bar{A}_k'\underline{V}_1\xi + \dots + \bar{C}_k'\underline{V}_1\zeta + S_{1k}'\underline{Y}_1\xi^2 + \dots\right\} \tag{3.93}$$

It wil be noticed that the <u>n</u>th order polynomial involves

$$(n + 1)\ (n + 2) \tag{3.94}$$

coefficients (<u>v</u>. §17d).

(d) In the paraxial limit $\underline{\Lambda}_j$ reduces to an optical invariant. In general we shall call any expression a <u>quasi-invariant</u> if it reduces to an optical invariant in the paraxial limit. And, as above, the change in a quasi-invariant in its passage across a refracting surface is of degree $\geqslant 3$ in the co-ordinates.

(e) It will be found convenient sometimes to write down all terms of the expansion (3.9). If the coeffieients of the terms of the <u>n</u>th order in $\Delta\underline{\Lambda}_j$ are $g^{(n)}_{\mu\nu,j}$, $\bar{g}^{(n)}_{\mu\nu,j}$ then we may write

$$\Delta\underline{\Lambda}_j = \sum_{n=1}^{\infty}\sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu}\left(g^{(n)}_{\mu\nu,j}\underline{Y}_{1} + \bar{g}^{(n)}_{\mu\nu,j}\underline{V}_{1}\right)\xi^{n-\mu}\eta^{\mu-\nu}\varsigma^{\nu} \tag{3.10}$$

Thus for example $g^{(1)}_{00,j} \equiv a_j$, ..., $\bar{g}^{(2)}_{21,j} \equiv \bar{s}_{5j}$; etc. Corresponding to (3.92) we have now

$$\left.\begin{aligned} G^{(n)}_{\mu\nu,j} &= \sum_{\alpha=1}^{j-1} g^{(n)}_{\mu\nu,\alpha} \\ \bar{G}^{(n)}_{\mu\nu,j} &= \sum_{\alpha=1}^{j-1} \bar{g}^{(n)}_{\mu\nu,\alpha}, \end{aligned}\right\} \tag{3.101}$$

whilst in the case of the dashed constants the summation is again to be extended to $\alpha = \underline{j}$ instead of to $\underline{j-1}$. (If $\underline{j} = 0$, the symbol $\sum_{\alpha=1}^{0}$ is to be interpreted as zero).

§4. Quasi-Linear Variables. Canonical Co-ordinates.

(a) In order to be able to carry out the expansions of §3.c. by a process of iteration (v. §5) it is necessary first to find explicit expressions for the departure from linearity of certain relations which have exactly linear equations as their paraxial counterparts. Let us briefly consider the latter.

A paraxial ray (y_1, u_1) has associated with it at any surface the quantities ℓ, y, u, i, ..., etc. Any pair of these may be used to specify the ray, provided they are independent. Then we shall call those quantities linear variables which are obtainable by a linear combination of y_1 and u_1, e.g. y_j, u_j, r_j, i_j, ... , or $au_i + bu_j$, $(a, b = const)$; etc.

If therefore α_j stands for any linear variable there will be a relation of the form

$$\alpha_j = \alpha_{pj}\, y_1 + \alpha_{qj}\, u_1 \tag{4.1}$$

in which the α_{pj}, α_{qj} are pure constants of the system. These are easily obtained in practice. It is only necessary to trace two (paraxial) rays, viz. the "p-ray" (1, 0), and the "q-ray" (0, 1). Then the actual value of α_j in the former trace will be α_{pj}; and the value of α_j in the latter will be α_{qj}.

We have the identity

$$\begin{vmatrix} y_{pj} & y_{qj} \\ u_{pj} & u_{qj} \end{vmatrix} = \frac{N_1}{N_j} \tag{4.11}$$

which follows at once from (3.21) for the special case of the rays (1, 0) and (0, 1). From (4.11) we find easily

$$\begin{vmatrix} \delta_{pj} & \delta_{qj} \\ u_{pj} & u_{qj} \end{vmatrix} = N_1 \quad ; \quad \text{and} \quad \begin{vmatrix} u_{pj} & u_{qj} \\ u'_{pj} & u'_{qj} \end{vmatrix} = \frac{N_1}{r_j} \Delta \frac{1}{N_j} = \varpi_j, \text{ say.} \tag{4.12}$$

In place of $\underline{y}_1$, $\underline{u}_1$ in (4.1) any other two independent linear variables might have been chosen, resulting in a different set of paraxial coefficients. But these can always be expressed in terms of the basic set above. (v. also (12.2)).

(b) Considering now non-paraxial rays, any variable which reduces to a linear variable in the paraxial limit will be called quasi-linear, or sometimes simply linear. If we choose four particular quasi-linear variables associated with a ray then these may be used to specify the latter; they are conveniently called the co-ordinates of the ray. The four variables of §3.c, i.e. Y_1, Z_1, V_1, W_1, are a particularly convenient set of co-ordinates; and we shall in future refer to them as canonical. (Linear co-ordinates in general are considered in §12). By definition we must have

$$\left.\begin{aligned} \underline{Y}_j &= y_{pj}\,\underline{Y}_1 + y_{qj}\,\underline{V}_1 + \underline{O}(3) \\ \underline{V}_j &= u_{pj}\,\underline{Y}_1 + u_{qj}\,\underline{V}_1 + \underline{O}(3) \end{aligned}\right\} \tag{4.2}$$

We now require explicit expressions for the remainder terms, which here appear merely as $\underline{O}(3)$. Without these, any attempt to calculate the higher order aberrations would be futile.

Replacing the axial ray in (3.71) by an arbitrary paraxial ray

$$\underline{\Lambda}_j = \underline{\Lambda}_1 + \sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_\nu \tag{4.21}$$

$$\left.\begin{aligned} \text{where now } \Lambda_y &= N\mu[(\ell - L)\tan U_y' + H] \\ \Lambda_z &= N\mu(\ell - L)\tan U_z \sec U_y \end{aligned}\right\} \tag{4.211}$$

$$\text{By definition} \quad \left.\begin{aligned} Y &= LV - H = L\tan U_y - H \\ Z &= LW = L\tan U_z \sec U_y \end{aligned}\right\} \tag{4.212}$$

so that (4.21) becomes

$$N_j(y_j \underline{V}_j - u_j \underline{Y}_j) = N_1(y_1 \underline{V}_1 - u_1 \underline{Y}_1) + \sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_\nu \tag{4.22}$$

Now any two different paraxial rays can be taken to be independent in the sense that all paraxial data are determined by just two such rays. Accordingly we choose the paraxial ray in (4.21) in two ways, viz. the p-ray (1, 0) and the q-ray (0, 1), leading to two independent equations of the form (4.22), i.e.

$$\left.\begin{aligned} N_j(y_{pj} \underline{V}_j - u_{pj} \underline{Y}_j) &= N_1 \underline{V}_1 + \sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{p\nu} \\ N_j(y_{qj} \underline{V}_j - u_{qj} \underline{Y}_j) &= -N_1 \underline{Y}_1 + \sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{q\nu} \end{aligned}\right\} \tag{4.221}$$

Solving these equations for $\underline{Y}_j$ and $\underline{V}_j$ we obtain, using (4.11)

$$\left.\begin{aligned} \underline{Y}_j &= y_{pj}(\underline{Y}_1 + \underline{\delta}_{yj}) + y_{qj}(\underline{V}_1 + \underline{\delta}_{vj}) \\ \underline{V}_j &= u_{pj}(\underline{Y}_1 + \underline{\delta}_{yj}) + u_{qj}(\underline{V}_1 + \underline{\delta}_{vj}) \end{aligned}\right\} \quad (4.3)$$

where we have written for the <u>increments</u>

$$\left.\begin{aligned} \underline{\delta}_{yj} &= -\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{q\nu} \\ \underline{\delta}_{vj} &= +\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{p\nu} \end{aligned}\right\} \quad (4.4)$$

(Here the first equation of (4.4) must be understood to represent the two equations

$\delta_{yj} = -\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\Lambda_{yq\nu}$, and $\delta_{zj} = -\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\Lambda_{zq\nu}$; and similarly in the case of the second equation).

The equations for the $\underline{Y}_j'$, $\underline{V}_j'$ are obtained by replacing $\underline{\delta}_{yj}$, $\underline{\delta}_{vj}$ by $\underline{\delta}_{yj}'$, $\underline{\delta}_{vj}'$ respectively, the latter differing from (4.4) only in that the summations are extended to $\nu = \underline{j}$ instead of to $\underline{j-1}$. (In the case of $\underline{V}_j$ u_{pj}, u_{qj} must of course also be replaced by u_{pj}', u_{qj}').

The increments in (4.4) are obviously $\underline{O}(3)$. (4.3) therefore represent the desired relations showing exactly the departure from linearity in the relations between the canonical variables. In (4.2) the $\underline{O}(3)$ of the first equation is simply

$$(y_{pj}\underline{\delta}_{yj} + y_{qj}\underline{\delta}_{vj}) = \delta\underline{Y}_j \text{ , say .} \qquad (4.41)$$

In the same way $\delta\underline{V}_j = (u_{pj}\underline{\delta}_{yj} + u_{qj}\underline{\delta}_{vj})$ (4.42)

The fact that the increments in the two equations of (4.3) are identical is of the utmost importance. For corresponding to any relation between (paraxial) linear variables we can now at once write down analogous equations holding between the corresponding canonical variables. Thus consider the equations

$$\left.\begin{aligned} i &= y/r - u \\ c &= N\,r\,i \end{aligned}\right\} \qquad (4.5)$$

If we define two angles I_y, I_z by

$$\sin \underline{I} = \underline{Y}/r - \underline{V} \qquad (4.51)$$

and hence

$$\underline{C} = Nr \sin \underline{I} \qquad (4.52)$$

then, by (4.3)

$$\underline{C}_j = c_{pj}(\underline{Y}_1 + \underline{\delta}_{yj}) + c_{qj}(\underline{V}_1 + \underline{\delta}_{vj}) \qquad (4.53)$$

Clearly, from (4.51), $r \sin I$ are the co-ordinates of the point in which the ray intersects the centre plane.

In Conrady's notation (§2)

$$\left.\begin{aligned} C_y &= N(L - r)\,(\tan U_y - \tan U_c) \\ C_z &= N(L - r)\,\tan U_z \sec U_y \end{aligned}\right\} \qquad (4.54)$$

Hence $\Delta(\alpha C_y) = \Delta C_y^* = \Delta N(L - r)\sin(U_y - U_c)\sec U_c \cos U_z$

$$= \Delta N(L - r)\sin U_z \cot\!an\, 0 \sec U_c$$

$$= 0 \;; \quad \text{by } (2.2\text{-}5)$$

and $$\Delta(\alpha C_z) = \Delta C_z^* = \Delta N(L - r)\sin U_z = 0$$

$$\therefore \Delta \underline{C}^* = 0 \qquad (4.55)$$

Also $$\frac{C_y}{C_z} = \cot 0 \sec U_c \qquad (4.56)$$

Accordingly (3.6) becomes $\Delta\Lambda_z = \dfrac{C_z}{C_y}\Delta\Lambda_y$

$$\text{or} \quad \frac{\Delta\Lambda_y}{C_y} = \frac{\Delta\Lambda_z}{C_z} = \underline{J}\text{, say.} \qquad (4.6)$$

(c) It will be seen that the canonical co-ordinates specify the ray in terms of that portion of it which is situated in the object space; nor do they in any way contain the positions of object, diaphragm or plane of reference. There are however other variables which have these valuable properties. Amongst these we single out a special set which is obtained by multiplying the canonical variables by α. In particular multiplying $\underline{Y}_1$, $\underline{V}_1$, by α_1, we obtain a set of <u>semi-canonical</u> co-ordinates, <u>viz</u>.

$$\left.\begin{aligned} \underline{Y}_1^* &= \alpha_1 \underline{Y}_1 \\ \underline{\beta}_1 &= \alpha_1 \underline{V}_1 \end{aligned}\right\} \qquad (4.7)$$

Corresponding to these we must now deal with the quasi-invariant $\underline{\Lambda}_j^* = \alpha_j \underline{\Lambda}_j$. Except in the case of $\underline{\beta}$ we shall throughout distinguish quantities referring explicitly or implicitly to semi-canonical co-ordinates by an asterisk; a notation which we have anticipated in (4.55). For example, we may at once write down

$$\underline{\beta}_j = u_{pj}(\underline{Y}_1^* + \underline{\delta}_{yj}^*) + u_{qj}(\underline{\beta}_1 + \underline{\delta}_{vj}^*) \qquad (4.8)$$

where

$$\left.\begin{aligned} \underline{\delta}^*_{yj} &= -\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}^*_{q\nu} \\ \underline{\delta}^*_{vj} &= +\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}^*_{p\nu} \end{aligned}\right\} \quad (4.81)$$

The sole advantage of these co-ordinates is that they allow of the direct calculation of direction cosines. As against this $\underline{H}^{*\prime}_k$ is not a very convenient quantity to deal with.

§5. Iteration.

(a) The equations developed in the last paragraph may now be employed to construct an iterative procedure whereby at each step we are able to calculate the exact aberration coefficients of next higher order. In this section we first consider the problem from the most general point of view.

At the jth surface $\Delta\underline{\Lambda}_j$ may be immediately expanded in terms of the canonical variables at that surface, viz.

$$\Delta\underline{\Lambda}_j = \sum_{n=1}^{\infty} \underline{\Gamma}_{n,j}(Y_j, Z_j, V_j, W_j) \quad , \qquad (5.1)$$

where $\Gamma_{n,j}$ is a homogeneous polynomial of degree 2n + 1. Let us replace the variables in (5.1) by the canonical co-ordinates by means of (4.3) omitting the increments. This leads directly to a "partial expansion" of the form

$$\Delta\underline{\Lambda}_j = \sum_{n=1}^{\infty} \underline{\Gamma}_{n,j}(Y_1, Z_1, V_1, W_1) \quad , \qquad (5.2)$$

where now $\underline{\Gamma}_{n,j}$ is a homogeneous polynomial of degree 2n + 1 in the co-ordinates. As in §3.e. this may be written in the general form

$$\underline{\Gamma}_{n,j} = \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (\underline{g}^{(n)}_{\mu\nu,j} \underline{Y}_{1} + \underline{\overline{g}}^{(n)}_{\mu\nu,j} \underline{V}_{1}) \xi^{n-\mu} \eta^{\mu-\nu} \varsigma^{\nu} \quad , \tag{5.21}$$

where the $\underline{g}^{(n)}_{\mu\nu,j}$, $\underline{\overline{g}}^{(n)}_{\mu\nu,j}$ are constants.

(As before we write for convenience in the case of the lower orders

$$\left.\begin{aligned} \underline{\Gamma}_{1,j} &= \underline{a}_j \underline{Y}_1 \xi + \underline{\overline{a}}_j \underline{V}_1 \xi + \dots \underline{\overline{c}}_j \underline{V}_1 \varsigma \; , \\ \underline{\Gamma}_{2,j} &= \underline{s}_{1j} \underline{Y}_1 \xi^2 + \dots \underline{\overline{s}}_{6j} \underline{V}_1 \varsigma^2 \; , \text{ etc.} \quad) \end{aligned}\right\} \tag{5.211}$$

Now let the <u>exact</u> expansion (<u>v</u>. eqs. (3.9), (3.10)) of $\Delta\underline{\Lambda}_j$ be

$$\Delta\underline{\Lambda}_j = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (g^{(n)}_{\mu\nu,j} \underline{Y}_1 + g^{(n)}_{\mu\nu,j} \underline{V}_1) \xi^{n-\mu} \eta^{\mu-\nu} \varsigma^{\nu} \tag{5.3}$$

Since the paraxial ray implicit in these equations is arbitrary (or in other words since the position of the object point is arbitrary) we must have

$$g^{(n)}_{\mu\nu,j} = g^{(n)}_{\mu\nu,j|p} y_{01} + g^{(n)}_{\mu\nu,j|q} u_{01} \tag{5.31}$$

and similarly for the other coefficients of the series (5.21) and (5.3). The $g^{(n)}_{\mu\nu,j|p}$, ... are then <u>pure</u> constants of the system.

By (4.4) and (3.101)

$$\left.\begin{aligned} \underline{\delta}_{yj} &= -\frac{1}{N_1} \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (G^{(n)}_{\mu\nu,j|q} \underline{Y}_1 + \overline{G}^{(n)}_{\mu\nu,j|q} \underline{V}_1) \xi^{n-\mu} \eta^{\mu-\nu} \varsigma^{\nu} \\ \underline{\delta}_{vj} &= +\frac{1}{N_1} \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} (G^{(n)}_{\mu\nu,j|p} \underline{Y}_1 + \overline{G}^{(n)}_{\mu\nu,j|p} \underline{V}_1) \xi^{n-\mu} \eta^{\mu-\nu} \varsigma^{\nu} \end{aligned}\right\} \tag{5.4}$$

Then, since by (4.3) the result of substituting $\underline{Y}_1 + \underline{\delta}_{yj}$, $\underline{V}_1 + \underline{\delta}_{vj}$ for $\underline{Y}_1$, $\underline{V}_1$ respectively in (5.2) must be to reproduce the series (5.3) we have the required identities

$$\sum_{n=1}^{\infty}\sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu}\Big\{\big[\underline{g}^{(n)}_{\mu\nu,j|p}(\underline{Y}_1 + \underline{\delta}_{yj}) + \bar{\underline{g}}^{(n)}_{\mu\nu,j|p}(\underline{V}_1 + \underline{\delta}_{vj})\big]\ (\xi+\delta\xi_j)^{n-\mu}(\eta+\delta\eta_j)^{\mu-\nu}(\varsigma+\delta\varsigma_j)^{\nu}$$

$$- (g^{(n)}_{\mu\nu,j|p}\ \underline{Y}_1 + \bar{g}^{(n)}_{\mu\nu,j|p}\underline{V}_1)\,\xi^{n-\mu}\eta^{\mu-\nu}\varsigma^{\nu}\Big\} \equiv 0 \quad , \qquad (5.5)$$

with a similar equation in which the $\underline{g}^{(n)}_{\mu\nu,j|p}$, ... are replaced by $\underline{g}^{(n)}_{\mu\nu,j|q}$, ... Here

$$\delta\xi_j \quad = \quad 2(Y_1\delta_{yj} + Z_1\ \delta_{zj}) + ((\delta_{yj})^2 + (\delta_{zj})^2), \text{ etc.} \qquad (5.51)$$

The coefficients of $\underline{Y}_1\,\xi^{n-\mu}\eta^{\mu-\nu}\varsigma^{\nu}$, $\underline{V}_1\,\xi^{n-\mu}\eta^{\mu-\nu}\varsigma^{\nu}$ in (5.5) must therefore vanish separately for all allowed values of n, μ, ν. The resulting equations make it possible to determine the required coefficients $g^{(n)}_{\mu\nu,j|p}$,... occurring in the expansions for the contributions without difficulty. For since the increments are $\underline{O}(3)$ the coefficients $G^{(m)}_{\mu\nu,j|p}$, ... occurring in the equations giving the $\underline{n}$th order coefficients $g^{(n)}_{\mu\nu,j|p}$, ... are of order less than $\underline{n}$ and are thus already known. In practice we therefore proceed as follows:

(b) The $\underline{g}^{(m)}_{\mu\nu,j|p}$, ... in (5.21) may be determined directly without difficulty ($\underline{v}$. §20). Then, by inspection of (5.5)

$$g^{(1)}_{\mu\nu,j|p} \quad = \quad \underline{g}^{(1)}_{\mu\nu,j|p}\ , \text{ etc.} \qquad (5.6)$$

from which the $G^{(1)}_{\mu\nu,j|p}$, ... follow directly from (3.101). The increments (5.4) are therefore now known correctly to the first order, all terms $\underline{O}(m)$, $(m>1)$ being omitted. Substituting these in (5.5), leaving out all terms $\underline{O}(m)$, $(m>2)$, the secondary coefficients $g^{(2)}_{\mu\nu,j|p}$, ... are obtained at once. By (5.4) the increments are then known correctly to the second order; and by substitution in (5.5), omitting all terms $\underline{O}(m)$, $(m>3)$, the tertiary coefficients follow directly. Obviously this process may be continued indefinitely, each step resulting in the knowledge of the coefficients of next higher order. Finally, the coefficients appearing in the required expansion of $\underline{\varepsilon_k}'$ are given by (3.101). (The summations are extended to $\alpha = \underline{k}$, and the final division by μ' is then carried out.) The secondary coefficients obtained by this method are given fully in §21.

(c) Since there are $\underline{4}$ paraxial coefficients per surface determining the intersection point of a ray with the image plane, the total number of coefficients determining the co-ordinates of this point correct to the $\underline{n}$th order is for a system of $\underline{k}$ refracting surfaces,

$$k\left\{4 + 2\sum_{m=1}^{n}(m+1)(m+2)\right\} = \tfrac{2}{3}k(n+1)(n+2)(n+3), \tag{5.7}$$

the position of the object being arbitrary (variable). If the position of the object is fixed this number may be reduced to $\frac{2}{3}k(n+1)(n+2)(n+\frac{3}{2})$. But far more important from a practical point of view is the fact that the calculation of the "$\underline{q}$-coefficients" is so similar to that of the "$\underline{p}$-coefficients" that the labour involved is virtually reduced to a half of what it would otherwise be. Although the coefficients are not all independent ($\underline{v}$. §17) the easiest course nevertheless is to calculate them independently. Many of the more complex terms can be arranged to recur frequently in the computing scheme so as further to minimise the work required. If semi-canonical co-ordinates are used the treatment is similar in all respects.

The development of specific algebraic expressions for $\Delta \underline{\Lambda}_j$, $\underline{J}$, etc. will be postponed until Part II. For the present we shall consider various problems which often demand our attention in practice.

§6. The Diaphragm.

(a) Let the diaphragm (i.e. the aperture stop) be situated a distance $\underline{d}$ behind the $\underline{i}$th surface, say. For convenience we shall take the rim of the aperture of the diaphragm to be circular and of radius ρ_i. Then the extreme rays of a pencil of rays from a given object point are frequently taken to be the rays grazing the rim of the diaphragm which is described in terms of polar co-ordinates ρ_i, θ_i by

$$\left.\begin{aligned} y &= \rho_i \cos \theta_i \\ z &= \rho_i \sin \theta_i \end{aligned}\right\} \quad (6.1)$$

$\underline{x}$, $\underline{y}$, $\underline{z}$ being the usual rectangular co-ordinate system in the space in which the diaphragm is situated. Hence, using paraxial relations, we have for extreme rays

$$\underline{Y}_i' - d\,\underline{V}_i' = \rho_i \begin{matrix}\cos\\ \sin\end{matrix}\,\theta_i \qquad (6.11)$$

Writing

$$y_{oi} - d\,u_{oi}' = \rho_o \qquad (6.12)$$

(6.11) gives,

$$\rho_o\,\underline{Y}_1 + \rho_q\,\underline{h}\,u_{o1} = \rho_i\,y_{o1} \begin{matrix}\cos\\ \sin\end{matrix}\,\theta_i \qquad (6.13)$$

since all rays pass through the object point, so that

$$\underline{V}_1 = (\underline{Y}_1 + \underline{h})/l_{01} \quad ; \quad (\underline{h} \equiv \underline{h}). \tag{6.131}$$

If the co-ordinates Y_1, Z_1 have the special values Y_o, 0 for the principal ray, and if ρ, θ are polar co-ordinates specifying extreme rays with respect to Y_o, 0, then

$$\left.\begin{aligned} Y_1 &= Y_o + \rho\cos\theta \\ Z_1 &= \rho\sin\theta \\ \text{and, by (6.131), } V_1 &= (Y_o + h + \rho\cos\theta)/l_{01} \\ W_1 &= \rho\sin\theta/l_{01} \end{aligned}\right\} \tag{6.2}$$

then (6.13) becomes

$$\left.\begin{aligned} \rho_o(Y_o + \rho\cos\theta) + \rho_q h u_{01} &\equiv \rho_i y_{01} \cos\theta_i \\ \rho_o\rho \sin\theta &\equiv \rho_i y_{01} \sin\theta_i \end{aligned}\right\} \tag{6.3}$$

$$\left.\begin{aligned} \therefore \quad \theta &= \theta_i \\ \rho &= \frac{\rho_i y_{01}}{\rho_o} \\ Y_o &= -\frac{\rho_q h u_{01}}{\rho_o} \end{aligned}\right\} \tag{6.4}$$

Also the paraxial location of the entrance pupil with respect to the first surface (v. T 10.1) is given by

$$p_1 = (Y_1/V_1)_{pr} = -\rho_q/\rho_p \tag{6.41}$$

whilst by (6.4), (6.41) the aperture of the

$$\text{entrance-pupil} = 2\rho_i/\rho_p \tag{6.42}$$

$\therefore$ Stop-number of system $= \rho_p/2\rho_i\, u'_{pk}$ (6.43)

Given the position of the diaphragm p_1, Y_o, etc. are therefore readily calculated.

(b) If the aberrations of the system be taken into account we obtain equations considerably more complex than (6.4); in particular the curve traced out by the extreme rays on the first tangent plane is no longer a circle. Equations (6.3) must now be replaced by

$$\left.\begin{aligned} \rho_o(Y_o + \rho\cos\theta) + \rho_q h u_{oi} + \left[\rho_p \delta'_{yi} + \rho_q \delta'_{vi}\right] &\equiv \rho_i y_{oi} \cos\theta_i \\ \rho_o \rho \sin\theta + \left[\rho_p \delta_{zi}' + \rho_q \delta_{wi}'\right] &\equiv \rho_i y_{oi} \sin\theta_i \end{aligned}\right\} (6.5)$$

where in the increments the co-ordinates Y_1, V_1 are replaced by Y_o, h, ρ, θ by means of (6.2). Putting $\rho = \rho_i = 0$ the first equation of (6.5) may be solved for Y_o; after which (14.4) may be solved as it stands for the remaining unknowns ρ and $\cos\theta$.

If aberrations are required only to the second order then Y_o, ρ, $\cos\theta$ are required only to the first order. These may be obtained directly by transposing the terms in square brackets in (6 .5) to the right hand side of these equations and then using in (6.2) the expression (6.4) for θ, ρ, Y_o.

(c) Problems in which the diaphragm is involved tend to become very laborious, partly on account of the complexity of the equations (6 .5), and partly because in a sense the term "entrance pupil" as generally understood loses its precise meaning when the optical system is not "perfect". Accordingly we shall use the convention that wherever the diaphragm may be situated we shall replace it by an <u>equivalent diaphragm</u> with an aperture of radius $\rho_i y_{o1}/\rho_o$ lying in the first tangent plane with centre $(-\rho_q h u_{o1}/\rho_o, 0)$. Principal and extreme rays are then defined in terms of the centre and the rim of the equivalent diaphragm. In practice this convention is scarcely likely to create any difficulties; whilst by means of it a consistent and much simplified theory can be built up.

§7. Shift of Image Plane.

It is sometimes necessary to consider the image produced by an optical system in a plane (perpendicular to the axis of symmetry) other than the ideal image plane. Accordingly let this out-of-focus plane intersect the axis in the point defined by $l_{ok}' + x'$. (The magnitude of the shift x' is not restricted). If $\bar{\underline{\varepsilon}}_k'$ is the displacement of the intersection point of the ray with this plane referred to the point $(l_{ok}' + x', h', 0)$ then clearly

$$\bar{\underline{\varepsilon}}_k' = \underline{\varepsilon}_k' + x'\underline{V}_k' \tag{7.1}$$

where $\underline{\varepsilon}_k'$ has its original meaning.

It may however be advantageous to define a new "out-of-focus ideal image point O_x'". Thus, by (6.2), $\underline{V}_k'$ may be written

30

$$\left.\begin{aligned} V_k' &= \frac{u_{ok}'}{y_{o1}}(Y_o + \rho\cos\theta) + \frac{u_{qk}'h}{l_{o1}} + \delta V_k' \\ W_k' &= \frac{u_{ok}'}{y_{o1}}\rho\sin\theta + \delta W_k' \end{aligned}\right\} \quad (7.2)$$

By (6.4)

$$\frac{u_{ok}'}{y_{o1}} Y_o + \frac{u_{qk}'h}{l_{o1}} = \frac{hu_{o1}(\rho_p u_{qk}' - \rho_q u_{pk}')}{\rho_o} \qquad (7.21)$$

Hence if the displacement $\bar{\varepsilon}_k'$ be now referred to the point O_x' with the co-ordinates

$$\left(l_{ok}' + x',\ h' + x'hu_{o1}(\rho_p u_{qk}' - \rho_q u_{pk}')/\rho_o\ ,\ 0\right) \qquad (7.22)$$

then $$\bar{\varepsilon}_k' = \varepsilon_k' + x'(\delta V_k' + \frac{u_{ok}'}{y_{o1}}\rho{\cos \atop \sin}\theta) \qquad (7.3)$$

In practice, when x' is sufficiently small, the term $x'\delta V_k'$ may often be omitted, so that we then simply have

$$\bar{\varepsilon}_k' = \varepsilon_k' + x'\frac{u_{ok}'}{y_{o1}}\rho{\cos \atop \sin}\theta \qquad (7.31)$$

§8. Shift of Diaphragm.

(a) From a practical point of view a change in the position of the diaphragm presents virtually no problem at all. Take ρ, θ as the independent variables specifying rays for a given position of the object. Since the $G^{(n)'}_{\mu\nu,k}/\mu'$ and the $\bar{G}^{(n)'}_{\mu\nu,k}/\mu'$ do not involve the position of the diaphragm at all we need only use the appropriate value of Y_o, as given by (6.4), in the equations (6.2) for any desired position of the diaphragm.

(b) A more theoretical study of this problem demands a precise answer to the question as to the particular aberration coefficients the consequent change of which is to be investigated. Clearly it is vain to consider the $G^{(n)'}_{\mu\nu,k}$, $\bar{G}^{(n)'}_{\mu\nu,k}$ for we already know that these are independent of the position of the diaphragm. When higher orders are considered the choice of different pairs of independent variables (such as polar co-ordinates in the planes of the paraxial entrance pupil or exit pupil respectively) will in some degree affect the results obtained; whereas in primary theory this is a matter of indifference. In virtue of our redefinition of the principal ray in terms of the centre of the equivalent diaphragm a suitable and simple choice is again the pair of polar co-ordinates ρ, θ (§6).

By means of (6.4) the terms of the expansion of $\underline{\varepsilon}_k{}'$ may then be rearranged in the form

$$\underline{\varepsilon}_k{}' = \sum_{n=1}^{\infty} \sum_{\mu=0}^{2n+1} \underline{\chi}^{(n)}_{\mu}(\theta)\rho^{2n+1-\mu} , \tag{8.1}$$

where the two sets of functions $\underline{\chi}^{(n)}_{\mu}$ depend on Y_o. Accordingly interest is centred around the change occurring in the coefficients in the expansion of $\underline{\chi}^{(n)}_{\mu}$ when Y_o assumes a new value. The change-over to (8.1) is facilitated by observing that the various powers of ξ,η,ς are expressible in the form (22.8). For instance if $\underline{s}$ is an integer

$$\begin{aligned} \xi^s &= (Y_o{}^2 + 2Y_o\rho\cos\theta + \rho^2)^s \\ &= Y_o{}^{2s}(1 + t + \tfrac{1}{4}\sec^2\theta t^2)^s \end{aligned} \tag{8.2}$$

$$\text{where} \quad t = \frac{2\cos\theta}{Y_o}\rho \tag{8.21}$$

The expansions of $\chi_{\mu}^{(n)}$ are of the form

$$\left.\begin{aligned} \chi_{\mu y}^{(n)} &= \sum_{\alpha} \sigma_{\mu,\alpha}^{(n)} \cos^{\alpha+1}\theta + \sum_{\beta} \tau_{\mu,\beta}^{(n)} \cos^{\beta}\theta \\ \chi_{\mu z}^{(n)} &= \sum_{\alpha} \sigma_{\mu,\alpha}^{(n)} \cos^{\alpha}\theta \sin\theta \end{aligned}\right\} \quad (8.3)$$

where the summations are extended over certain integral values of α and β.

We are thus concerned with the changes $\delta\sigma_{\mu,\alpha}^{(n)}$, $\delta\tau_{\mu,\beta}^{(n)}$ which occur in $\sigma_{\mu,\alpha}^{(n)}$, $\tau_{\mu,\beta}^{(n)}$ when Y_0 is changed by δY_0, <u>where the latter need not be restricted in magnitude.</u>

(c) As in the case of the well-known results for the primary aberration coefficients (<u>v.</u> <u>Conrady</u>, 1929 (c)) the $\delta\sigma_{\mu,\alpha}^{(n)}$, $\delta\tau_{\mu,\beta}^{(n)}$ ($\mu = 0, 1, 2, \ldots$) may then be expressed in terms of the $\sigma_{\nu,\alpha}^{(n)}$, $\tau_{\nu,\beta}^{(n)}$ ($\nu = \mu - 1, \mu - 2, \ldots$) themselves. As a general discussion would lead too far here we content ourselves with merely quoting the required result for the first few coefficients (<u>i.e.</u> for $\mu = 0, 1, 2$), <u>viz.</u>

$$\left.\begin{aligned} \delta\sigma_{0,0}^{(n)} &= 0 \\ \delta\sigma_{1,1}^{(n)} &= 2n\,\sigma_{0,0}^{(n)} \cdot \delta Y_0 \\ \delta\sigma_{2,0}^{(n)} &= \sigma_{1,1}^{(n)}\,\delta Y_0 + n\sigma_{0,0}^{(n)}(\delta Y_0)^2 \\ \delta\sigma_{2,2}^{(n)} &= 2(n-1)\,\delta\sigma_{2,0}^{(n)} \\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ \delta\tau_{1,0}^{(n)} &= 0 \\ \delta\tau_{2,1}^{(n)} &= 2n\tau_{1,0}^{(n)} \cdot \delta Y_0 \end{aligned}\right\} \quad (8.4)$$

These equations are briefly discussed in §13.c.

§9. Change in Position of Object.

(a) As in §8 we first consider changes in object position from a purely practical standpoint. Let the object be moved along the axis through a distance $\underline{x}$. We then wish to determine the aberration $\bar{\underline{\varepsilon}}_k'$ of the ray with respect to the new ideal image point. This has of course the co-ordinates $(l_{ok}' + x', \bar{h}', 0)$, where the ideal image height is

$$\bar{h}' = \frac{\mu}{\bar{\mu}'} h = \frac{N_1 h}{N_k'[u_{pk}'(l_{o1}+x) + u_{qk}']} \tag{9.1}$$

and the shift of the image plane is

$$x' = \frac{N_k'\mu^2 x}{N_1 \mu' \bar{\mu}'} \tag{9.11}$$

Since the $G^{(n)'}_{\mu\nu,k|p}$, ... do not involve l_{o1} we have

$$\begin{aligned}
\bar{\underline{\varepsilon}}_k' - \underline{\varepsilon}_k' &= \frac{1}{\bar{\mu}'} \bar{\underline{D}}_k' - \frac{1}{\mu} \underline{D}_k' \\
&= \frac{N_1}{N_k'} \left\{ \frac{\underline{\delta}_{vk}'(l_{o1}+x) - \underline{\delta}_{yk}'}{u_{pk}'(l_{o1}+x)+u_{qk}'} - \frac{\underline{\delta}_{vk}' l_{o1} - \underline{\delta}_{yk}'}{u_{pk}' l_{o1} + u_{qk}'} \right\} \\
&= \frac{N_k'}{N_1} \frac{\mu^2 x}{\mu' \bar{\mu}'} (u_{qk}' \underline{\delta}_{vk}' + u_{pk}' \underline{\delta}_{yk}')
\end{aligned}$$

∴ By (4.42), (9.1)

$$\bar{\underline{\varepsilon}}_k' = \underline{\varepsilon}_k' + x'\delta \underline{V}_k' \tag{9.2}$$

On the right-hand side of the equation the canonical co-ordinates are now related by

$$\underline{V}_1 = (\underline{Y}_1 + \underline{h})/(l_{o1} + x) \tag{9.21}$$

(9.2) and (9.21) is all that is necessary to deal in practice with unrestricted changes in the position of the object.

(b) If the image plane is not shifted we need only apply the results of §7 as well; after the change above has been made a shift through $-\underline{x}'$ restores the image plane to its original position. Using the exact equation (7.3) we therefore obtain the result (the magnitude of $\underline{x}$ is not restricted!)

$$\tilde{\underline{\varepsilon}}_k' = \underline{\varepsilon}_k' - \frac{m'x}{l_{01} + x}\,\rho \begin{matrix}\cos\\ \sin\end{matrix}\,\theta \tag{9.3}$$

where $\underline{\varepsilon}_k'$ implies (9.21) and (9.4), and $\tilde{\underline{\varepsilon}}_k'$ refers to the "new ideal image point" O_x'.

(c) In a more theoretical treatment of the effects of change of object position the remarks of §8.b concerning the choice of independent variables also apply. If in conformity with usual practice in primary theory we again choose the polar co-ordinates ρ, θ we shall have expansions of the form (8.1) and (8.3). But on account of the complex transition from (6.2) to (9.21) where at the same time the equation

$$Y_1 = Y_o + \rho \cos \theta$$

is replaced by

$$Y_1 = \overline{Y}_o + \rho\cos\theta = \frac{Y_o}{1 + \frac{\rho_p u_{\alpha 1} x}{\rho_o}} + \rho \cos\theta \tag{9.4}$$

the explicit expressions for the changes in these aberration coefficients become rather involved. Accordingly lack of space does not permit us here further to consider this problem.

§10. Reversed Optical Systems.

As usual let an object point $O(l_{01}, h, 0)$ be given possessing a conjugate ideal image point $O'(l_{ok}', m'h, 0)$, and let the aberration $\underline{\varepsilon}_k'$ be known explicitly as a function of the four canonical variables $\underline{Y}_1, \underline{V}_1$ (which are connected in pairs by the additional relations (6.131). Now let a pencil of rays issue from O'. Then a one-to-one correspondence can be established between the rays of the pencil from O' and that from O. This correspondence is to some extent arbitrary, except that we naturally demand that corresponding rays should pursue paths through the system which (in general) differ little from one another. Accordingly let us agree to the convention that corresponding rays shall be parallel in the space in which O lies. Let the aberration of the ray $\underline{R}'$ from O' be $\underline{\varepsilon}$ in the space in which O lies and let $\underline{\varepsilon}'(= \underline{\varepsilon}_k')$ have its usual meaning. If we distinguish the canonical co-ordinates of $\underline{R}'$ by a bar

$$\underline{\varepsilon} = l_{01}\underline{\overline{V}}_1 - \underline{\overline{Y}}_1 - \underline{h} \quad , \tag{10.1}$$

where $\overline{Y}_1$ and $\overline{V}_1$ are related by the equation

$$\underline{\overline{H}}_k' \equiv m'\underline{h} \quad . \tag{10.2}$$

By (10.1) and (3.71) the last equation may be written

$$\underline{\overline{H}}_k' = m'(l_{01}\underline{\overline{V}}_1 - \underline{\overline{Y}}_1 - \underline{\varepsilon})$$

$$= m'(l_{01}\underline{\overline{V}}_1 - \overline{Y}_1) + \underline{\varepsilon}'(\overline{Y}_1, \overline{Z}_1, \overline{V}_1, \overline{W}_1) \tag{10.21}$$

But according to our convention above

$$\underline{\bar{V}}_1 = \underline{V}_1 \quad . \tag{10.3}$$

Hence $\quad \underline{\bar{Y}}_1 = l_{o1}\underline{\bar{V}}_1 - \underline{h} - \underline{\varepsilon} = l_{o1}\underline{V}_1 - \underline{h} - \underline{\varepsilon}$.

$$\therefore \quad \underline{\bar{Y}}_1 = \underline{Y}_1 - \underline{\varepsilon} \quad . \tag{10.31}$$

Using (10.3), (10.31) in (10.21) we find

$$\left.\begin{aligned} \varepsilon_y'(Y_1 - \varepsilon_y, Z_1 - \varepsilon_z, V_1, W_1) + m'\varepsilon_y &= 0 \quad , \\ \varepsilon_z'(Y_1 - \varepsilon_y, Z_1 - \varepsilon_z, V_1, W_1) + m'\varepsilon_z &= 0 \quad . \end{aligned}\right\} \tag{10.4}$$

This is a pair of simultaneous algebraic equations with ε_y, ε_z as unknowns.

When terms of the third degree alone are taken into account we have at once

$$\underline{\varepsilon}' + m'\underline{\varepsilon} = 0 \quad . \tag{10.5}$$

This last equation constitutes an extension to extra-axial primary aberrations of a theorem on primary spherical aberration given by Conrady (Conrady, 1929 (d)). Moreover the form of the equations (10.4) suggests immediately that (10.5) may be a good approximation to the true result when higher orders are taken into account.

§11. Addition of Optical Systems.

(a) Suppose all computations have been carried out separately for two optical systems $\underline{S}$ and $\underline{\hat{S}}$. Then it may be of interest to know how the aberration coefficients of the composite system $\underline{S}^+$ are composed of those of $\underline{S}$ and $\hat{\underline{S}}$. Let $\underline{S}$ and $\hat{\underline{S}}$ be separated by a distance $\underline{s}$ measured along their common axis of symmetry; and without loss of generality we may suppose $\hat{\underline{S}}$ to lie between the object and $\underline{S}$.

It is sufficient here to indicate how the primary coefficients may be dealt with since the procedure is essentially the same for higher orders. Clearly, if $(\underline{D}')^+$ refers to $\underline{S}^+$

$$(\underline{D}')^+ = \hat{\underline{D}}_k' + \underline{D}_k' \tag{11.1}$$

where the suffix $\underline{k}$ in each case refers to the system in question.

To simplify the notation we shall write <u>in this paragraph</u>

$$\left.\begin{aligned} \hat{y}_k - s\hat{u}_k' &= z \\ \hat{u}_k' &= v \end{aligned}\right\} \tag{11.2}$$

$$\text{Then} \quad \underline{Y}_1 = \hat{\underline{Y}}_k' - s\hat{\underline{V}}_k' = z_p(\hat{\underline{Y}}_1 + \hat{\underline{\delta}}_{yk}') + z_q(\hat{\underline{V}}_1 + \hat{\underline{\delta}}_{vk}')$$

$$\underline{V}_1 = \hat{\underline{V}}_k' \tag{11.3}$$

From (11.3) ξ, η, ς may be calculated in terms of $\hat{\xi}, \hat{\eta}, \hat{\varsigma}$, thus

$$\xi = z_p{}^2\hat{\xi} + 2z_pz_q\hat{\eta} + z_q{}^2\hat{\varsigma} + \underline{O}(4) \text{ ; etc.} \tag{11.31}$$

$$\text{Hence} \quad \underline{D}_k' = A_k'(z_p\hat{\underline{Y}}_1 + z_q\hat{\underline{V}}_1)(z_p{}^2\hat{\xi} + 2z_pz_q\hat{\eta} + z_q{}^2\hat{\varsigma}) + \ldots$$
$$+ \overline{C}_k'(v_p\hat{\underline{Y}}_1 + v_q\hat{\underline{V}}_1)(v_p{}^2\hat{\xi} + 2v_pv_q\hat{\eta} + v_q{}^2\hat{\varsigma}) + \underline{O}(5) \tag{11.4}$$

Substituting in (11.1) we find without difficulty, omitting the suffix $\underline{k}$

$$\left.\begin{aligned}(A')^+ &= \hat{A}' + \{z_p{}^3A' + z_p{}^2v_p(\bar{A}' + B') + z_pv_p{}^2(\bar{B}' + C') + v_p{}^3\bar{C}'\}\\ (\bar{A}')^+ &= \hat{\bar{A}}' + \{z_p{}^2z_qA' + z_p{}^2v_q\bar{A}' + z_pz_qv_pB' + z_pv_pv_q\bar{B}' + z_qv_p{}^2C' + v_p{}^2v_q\bar{C}'\}\end{aligned}\right\} \quad (11.5)$$

. . . .

Now $\hat{A}' = A_p{}'\hat{y}_{01} + A_q{}'\hat{u}_{01}$

and $A' = A_p'y_{01} + A_q'u_{01} = (z_pA_p' + v_pA_q')\hat{y}_{01} + (z_qA_p' + v_qA_q')\hat{u}_{01}$ (11.51)

Hence we find for the $\underline{\text{pure}}$ constants of the compound system $\underline{S}^+$

$$\begin{aligned}(A_p')^+ = \hat{A}_p' &+ \{z_p{}^4A_p' + z_p{}^3v_p(A_q' + \bar{A}_p' + B_p') + z_p{}^2v_p{}^2(\bar{A}_q' + B_q' + \bar{B}_p' + C_p')\\ &+ z_pv_p{}^3(\bar{B}_q' + C_q' + \bar{C}_p') + v_p{}^4\bar{C}_q'\} ,\end{aligned} \quad (11.6)$$

and similarly for all other coefficients.

In calculating $(\varepsilon_k')^+$ it must be remembered that $(\mu')^+$ depends (linearly) on $\underline{s}$.

(b) It may be easily verified by direct calculation that the identities between the aberration coefficients (§17) for $\underline{S}^+$ continue to hold of course. But it is of additional interest to notice that we can show in this way that

$$(C_p')^+ - \tfrac{1}{2}(\bar{B}_p')^+ = (\hat{C}_p' - \tfrac{1}{2}\hat{\bar{B}}_p') + (\hat{N}_1/\hat{N}_k')^2(C_p' - \tfrac{1}{2}\bar{B}_p') \quad (11.7)$$

But the right-hand side of (11.7) is quite independent of $\underline{s}$; hence we have shown that there exists a combination of coefficients which is independent of the separation of the constituent surfaces of the system without knowing explicitly how this expression is made up of the radii, etc. of the system. Reference to (17.15) will show that we have in fact proved a well-known property of the Petzval Curvature.

(c) The distance $\underline{s}$ enters into the primary coefficients of $(\underline{\Lambda}'_k)^+$ only through z_p and z_q; and it will be seen from (11.6) that they therefore express themselves as polynomials of the fourth degree in $\underline{s}$. In the same way the coefficients of $(\underline{\Lambda}'_k)^+$ of order $\underline{n}$ will be polynomials of order $\underline{2n+2}$ in $\underline{s}$. (The remark made at the end of section (a) should be kept in mind here).

Since every optical system can be thought of as being made up of such component systems the results above afford a method (in principle at least) of investigating the exact changes in the aberrations which occur when the distances between the surfaces of the system are altered.

§12. <u>General Linear Co-ordinates</u>.

In practice it may sometimes be convenient to use special co-ordinates appropriate to the particular problem in hand. For instance in a given case the co-ordinates of the intersection points of the ray with the object plane and with the plane of the paraxial entrance pupil respectively might be quantities more suitable for the specifications of the ray than the canonical co-ordinates (although the aberration coefficients will then no longer be pure constants of the system).

We therefore define a more general set of linear co-ordinates

S_y, T_y, S_z, T_z by the equations

$$\left.\begin{aligned} \underline{S} &= \pi \underline{Y}_1 + \rho \underline{V}_1 + \sigma \underline{Y}_k' + \tau \underline{V}_k' \\ \underline{T} &= \bar{\pi} \underline{Y}_1 + \bar{\rho} \underline{V}_1 + \bar{\sigma} \underline{Y}_k' + \bar{\tau} \underline{V}_k' \quad , \end{aligned}\right\} \quad (12.1)$$

where in this paragraph $\pi, \ldots, \bar{\tau}$ are given constants. (It is of course conceivable that some intermediate co-ordinates, such as $\underline{Y}_j$, $\underline{V}_j$ $(1<j\leq k)$ might be convenient; but this is not likely often to be the case. If necessary, equations similar to those below may however be developed by the same methods).

A paraxial linear variable α_j is now related to $\underline{s}$ and $\underline{t}$ by an equation of the form

$$\alpha_j = \alpha_{aj} s + \alpha_{bj} t \quad , \qquad (12.2)$$

where

$$\left.\begin{aligned} s &= \pi y_1 + \rho u_1 + \sigma y_k + \tau u_k' \\ t &= \bar{\pi} y_1 + \bar{\rho} u_1 + \bar{\sigma} y_k + \bar{\tau} u_k' \quad , \end{aligned}\right\} \quad (12.21)$$

and the α_{aj}, α_{bj} are constants.

Substituting (12.21) in (12.2) and using (4.1) it follows that

$$\left.\begin{aligned} \alpha_{pj} &= (\pi + \sigma y_{pk} + \tau u_{pk}')\alpha_{aj} + (\bar{\pi} + \bar{\sigma} y_{pk} + \bar{\tau} u_{pk}')\alpha_{bj} \\ \alpha_{qj} &= (\rho + \sigma y_{qk} + \tau u_{qk}')\alpha_{aj} + (\bar{\rho} + \bar{\sigma} y_{qk} + \bar{\tau} u_{qk}')\alpha_{bj} \quad , \end{aligned}\right\} \quad (12.22)$$

which we shall write

$$\left.\begin{aligned} \alpha_{pj} &= g_p \alpha_{aj} + \bar{g}_p \alpha_{bj} \\ \alpha_{qj} &= g_q \alpha_{aj} + \bar{g}_q \alpha_{bj} \quad . \end{aligned}\right\} \quad (12.23)$$

Clearly the set of constants $\pi, \ldots, \bar{\tau}$ must be chosen such that

$$g = \begin{vmatrix} g_p & \bar{g}_p \\ g_q & \bar{g}_q \end{vmatrix} \neq 0 \quad . \tag{12.231}$$

(If $g = 0$ it follows that $\underline{s}$ and $\underline{t}$ are not independent).

We then have

$$\begin{vmatrix} y_{aj} & y_{bj} \\ u_{aj} & u_{bj} \end{vmatrix} = N_1/N_j g \quad . \tag{12.232}$$

(12.21) may now be written

$$\left.\begin{aligned} s &= g_p y_1 + g_q u_1 \\ t &= \bar{g}_p y_1 + \bar{g}_q u_1 \quad . \end{aligned}\right\} \tag{12.24}$$

Hence if the $\underline{a}$ and $\underline{b}$ rays be the rays $s = 1$, $t = 0$ and $s = 0$, $t = 1$ respectively, it follows at once that

$$\left.\begin{aligned} \underline{a}\text{-ray} &= (+\bar{g}_q/g, -\bar{g}_p/g) \\ \underline{b}\text{-ray} &= (-g_q/g, +g_p/g) \quad . \end{aligned}\right\} \tag{12.25}$$

Any numerical work accordingly now begins with the tracing of these two rays.

(b) The equations (4.221) may be transposed so as to express $\underline{Y}_1$, $\underline{V}_1$ in terms of $\underline{Y}_j$, $\underline{V}_j$. In the same way the equation

$$\underline{\Lambda}_j = \underline{\Lambda}_k' - \sum_{\nu=j}^{k} \Delta\underline{\Lambda}_\nu \tag{12.3}$$

when written in component form may be solved for $\underline{Y}_k'$, $\underline{V}_k'$ in terms of $\underline{Y}_j$, $\underline{V}_j$. By means of (12.23) the $\Delta\underline{\Lambda}_{p\nu}$, $\Delta\underline{\Lambda}_{q\nu}$

72

may be expressed in terms of $\Delta\underline{\Lambda}_{a\nu}$, $\Delta\underline{\Lambda}_{b\nu}$. Substituting the expressions for $\underline{Y}_1$, $\underline{V}_1$, $\underline{Y}_k'$, $\underline{V}_k'$ so obtained in (12.1), we find after some rearranging of terms, and using (12.23),

$$\left.\begin{aligned} \underline{Y}_j &= y_{aj}(\underline{S} + \underline{\delta}_{sj}) + y_{bj}(\underline{T} + \underline{\delta}_{tj}) \\ \underline{V}_j &= u_{aj}(\underline{S} + \underline{\delta}_{sj}) + u_{bj}(\underline{T} + \underline{\delta}_{tj}) \quad , \end{aligned}\right\} \quad (12.4)$$

where

$$\left.\begin{aligned} \underline{\delta}_{sj} &= \frac{1}{N_1}\left\{+g \sum_{\nu=j}^{k} \Delta\underline{\Lambda}_{b\nu} + \begin{vmatrix} g_p & g_q \\ \pi & \rho \end{vmatrix} \sum_{\nu=1}^{k} \Delta\underline{\Lambda}_{a\nu} + \begin{vmatrix} \bar{g}_p & \bar{g}_q \\ \pi & \rho \end{vmatrix} \sum_{\nu=1}^{k} \Delta\underline{\Lambda}_{b\nu}\right\} \\ \underline{\delta}_{tj} &= \frac{1}{N_1}\left\{-g \sum_{\nu=j}^{k} \Delta\underline{\Lambda}_{a\nu} + \begin{vmatrix} g_p & g_q \\ \bar{\pi} & \bar{\rho} \end{vmatrix} \sum_{\nu=1}^{k} \Delta\underline{\Lambda}_{a\nu} + \begin{vmatrix} \bar{g}_p & \bar{g}_q \\ \bar{\pi} & \bar{\rho} \end{vmatrix} \sum_{\nu=1}^{k} \Delta\underline{\Lambda}_{b\nu}\right\} . \end{aligned}\right\} (12.41)$$

The $\Delta\underline{\Lambda}$ occurring in (12.41) are of course now calculated by means of (12.4). The expressions for $\underline{\delta}_{sj}'$,$\underline{\delta}_{tj}'$ are obtained from (12.41) by merely replacing the summations over ν from $\underline{j}$ to $\underline{k}$ by summations from $\underline{j+1}$ to $\underline{k}$; and in every case we use the convention

$$\sum_{\nu=1}^{0} \equiv 0 \; ; \qquad \sum_{\nu=k+1}^{k} \equiv 0 \quad . \qquad (12.411)$$

As a simple example we shall consider the case mentioned at the beginning of section (a). Thus if $\underline{p}_1$ is the paraxial location location of the centre of the entrance pupil ($\underline{v}.$ (6.41)) we have

$$\left.\begin{aligned} \pi &= -1 & \bar{\pi} &= -1 \\ \rho &= p_1 & \bar{\rho} &= l_{01} \\ \sigma &= 0 & \bar{\sigma} &= 0 \\ \tau &= 0 & \bar{\tau} &= 0 \; . \end{aligned}\right\} (12.5)$$

$$\left.\begin{aligned} \therefore\; g_p = -1; \quad \bar{g}_p = -1 \; ; \quad g_q = p_1 \; ; \quad \bar{g}_q = l_{01} \\ \text{and} \quad g = p_1 - l_{01} \; . \end{aligned}\right\} (12.51)$$

The condition (12.231) therefore merely requires that the object shall not lie in the plane of the paraxial entrance pupil. In virtue of (12.5), (12.51) the equations (12.41) reduce to

$$\left.\begin{aligned} \underline{\delta}_{sj} &= -\frac{g}{N_1}\sum_{\nu=1}^{j-1}\Delta\underline{\Lambda}_{b\nu} \\ \underline{\delta}_{tj} &= +\frac{g}{N_1}\sum_{\nu=1}^{j-1}\Delta\underline{\Lambda}_{a\nu} \quad , \end{aligned}\right\} \quad (12.52)$$

which may be compared with (4.4).

All the equations (12.3) to (12.52) have their analogous counter-parts when semi-canonical variables replace those appearing in (12.1). Thus, in the case of the mixed angle-variables

$$\left.\begin{aligned} \underline{S}^* &= \underline{\beta}_1 \\ \underline{T}^* &= \underline{\beta}_k , \end{aligned}\right\} \quad (12.6)$$

we find

$$\left.\begin{aligned} \underline{\delta}^*_{sj} &= +\frac{1}{N_1 f'}\sum_{\nu=1}^{j-1}\Delta\underline{\Lambda}^*_{b\nu} \\ \underline{\delta}^*_{tj} &= +\frac{1}{N_1 f'}\sum_{\nu=j}^{k}\Delta\underline{\Lambda}^*_{a\nu} \quad . \end{aligned}\right\} \quad (12.61)$$

(c) Whatever the particular set of constants $\pi, \ldots, \bar{\tau}$ may be, we shall write in the case of the general linear co-ordinates

$$\Delta\underline{\Lambda} = \sum_{n=1}^{\infty}\sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu}\left(f^{(n)}_{\mu\nu,j}\,\underline{S} + \bar{f}^{(n)}_{\mu\nu,j}\,\underline{T}\right)\tilde{\xi}^{\,n-\mu}\tilde{\eta}^{\,\mu-\nu}\tilde{\zeta}^{\,\nu} \quad , \qquad (12.7)$$

where

$$\left.\begin{aligned} \tilde{\xi} &= S_y^{\,2} + S_z^{\,2} \\ \tilde{\eta} &= S_yT_y + S_zT_z \\ \tilde{\zeta} &= T_y^{\,2} + T_z^{\,2} \end{aligned}\right\} \quad (12.71)$$

corresponding to (3.10); and similarly

$$\sum_{\alpha=1}^{j-1} f^{(n)}_{\mu\nu,\alpha} = \quad F^{(n)}_{\mu\nu,j}\text{; etc.} \tag{12.72}$$

It is also possible to convert the series (3.10) into (12.7) and so to obtain the two sets of coefficients in terms of one another. Using (4.3) in (12.1), and inverting the resulting series, $\underline{Y}_1$ and $\underline{V}_1$ can be expressed as functions of $\underline{S}$, $\underline{T}$, the coefficients in the series being those appearing in (3.10). Substituting these in (3.10) we must of necessity obtain (12.7); and by comparing coefficients the $f^{(n)}_{\mu\nu,j}$, ... can be written down in terms of the $g^{(n)}_{\mu\nu,j|p}$, ... In general this procedure is so laborious as to make the use of the more direct methods outlined above advisable.

§13. On the Classification of Aberrations.

(a) The higher order aberrations may be classified in different ways according to the type of problem under consideration. Thus we may define certain aberrations in a purely geometrical way, as dictated by convenience, and then examine these in relation to the coefficients discussed here (v. T§3.b). Or else we may expand $\underline{\varepsilon}_k{}'$ in terms of some convenient quantity such as $\underline{h}$ and then consider the properties of the different terms of the resulting series (Steward, 1926). For this purpose it is useful to employ the co-ordinates given by (12.5). Then

$$\mu'\underline{\varepsilon}_k{}' = \sum_{n=1}^{\infty}\sum_{\mu=0}^{n}\sum_{\nu=0}^{\mu}\left(F^{(n)'}_{\mu\nu,k}\,\rho\,{\cos \atop \sin}\,\theta + \overline{F}^{(n)'}_{\mu\nu,k}\,{h \atop 0}\right)\rho^{2n-(\mu+\nu)}h^{\mu+\nu}\cos^{\mu-\nu}\theta \tag{13.1}$$

We notice incidentally that the expansions according to powers of ρ and of $\underline{h}$ are materially the same.

(b) A general discussion would lead too far here. To illustrate the examination of our <u>known</u> coefficients we content ourselves with the case in which $\underline{h}$ is sufficiently small. To simplify the notation let us write in this paragraph

$$\left.\begin{aligned} \frac{1}{\mu'} F^{(n)\prime}_{00,k} &= \alpha_n ; \quad \frac{1}{\mu'} \overline{F}^{(n)\prime}_{00,k} = \overline{\alpha}_n \\ \frac{1}{2\mu'} F^{(n)\prime}_{10,k} &= \beta_n \quad . \end{aligned}\right\} \qquad (13.11)$$

Then the terms independent of $\underline{h}$ are

$$(\underline{\varepsilon}_k{}')_{h=0} = \left(\sum_{n=1}^{\infty} \alpha_n \rho^{2n+1}\right) \begin{matrix}\cos \\ \sin\end{matrix}\, \theta \quad . \qquad (13.2)$$

Thus α_n is the $\underline{n}$th order <u>spherical aberration</u> coefficient. For the terms linear in $\underline{h}$ we find

$$\left.\begin{aligned} (\varepsilon_{yk}{}')_1 &= \sum_{n=1}^{\infty} (\overline{\alpha}_n + 2\beta_n \cos^2\theta)\rho^{2n} h \\ (\varepsilon_{zk}{}')_1 &= \sum_{n=1}^{\infty} 2\beta_n \cos\theta \sin\theta\, \rho^{2n}\, h \quad . \end{aligned}\right\} \qquad (13.3)$$

If we isolate the terms of order $\underline{n}$ we have

$$\left.\begin{aligned} (\varepsilon_{yk}{}')_1^{(n)} &= \left\{\beta_n \cos 2\theta + (\overline{\alpha}_n + \beta_n)\right\}\rho^{2n} h \\ (\varepsilon_{zk}{}')_1^{(n)} &= \left\{\beta_n \sin 2\theta\right\} \rho^{2n} h \quad . \end{aligned}\right\} \qquad (13.31)$$

These represent a family of circles of radius $\beta_n \rho^{2n} h$, and centres $((\overline{\alpha}_n + \beta_n)\rho^{2n} h, 0)$, using a local co-ordinate system with origin at O'. The terms (13.31) are therefore said to characterise circular coma of order n. For varying ρ these circles have as envelope a pair of straight lines through O' making an angle

$$\psi_n = 2 \operatorname{arccosec}(1 + \bar{\alpha}_n/\beta_n) \tag{13.4}$$

with one another. Using a result to be derived later (§17e) a certain combination of the two coefficients occurring in (13.4) can be expressed in terms of the coefficients of order less than n. If the latter vanish identically we have in particular (for this position of the object)

$$\beta_n = n\bar{\alpha}_n \tag{13.41}$$

by (17.71) and §17.e.

$$\therefore \quad \psi_n = 2 \operatorname{arc\,cosec}(1 + \frac{1}{n}) \quad . \tag{13.42}$$

Equations (13.41) and (13.42) are mostly of academic interest. However, in general in the absence of spherical aberration coma is the outstanding aberration for small h. Considering only primary and secondary terms we may make use of the identity (17.93) in (13.31). Then remembering that we have assumed $\alpha_1 = 0$ this may be simplified to give

$$\beta_2 - 2\bar{\alpha}_2 = \frac{2}{m'} \beta_1{}^2 + \frac{l_{\alpha 1}{}^2}{g^2} \beta_1 \quad . \tag{13.5}$$

Hence, from (13.31)

$$\left.\begin{aligned} (\varepsilon'_{yk})_1^{(2)} &= \Big\{\beta_1(2 + \cos 2\theta)\rho^2 + \big[\beta_2 \cos 2\theta + \tfrac{3}{2}\beta_2 - \beta_1{}^2/m' \\ &\qquad - l_{\alpha 1}{}^2\beta_1/2g^2)\big]\,\rho^4\Big\}h \\ (\varepsilon'_{zk})_1^{(2)} &= \Big\{\beta_1 \sin 2\theta\, \rho^2 + \beta_2 \sin 2\theta\, \rho^4\Big\}h \,. \end{aligned}\right\} \tag{13.51}$$

The second order terms alone give a family of circles as above. But now

$$\psi_2 = 2 \text{ arccosec } \left(\frac{3}{2} - \frac{\beta_1}{\beta_2}\left[\frac{\beta_1}{m'} + \frac{l_{01}^{\ 2}}{2g^2}\right]\right) \tag{13.52}$$

which may be contrasted with (13.42). Clearly the coefficients β_1 and β_2 completely govern primary and secondary circular coma.

(c) Referring now to §8.b we must have

$$\left.\begin{aligned} \sigma_{0,0}^{(n)} &= \alpha_n \\ \sigma_{1,1}^{(n)} &= 2\beta_n h \\ \tau_{1,0}^{(n)} &= \bar{\alpha}_n h \end{aligned}\right\} \tag{13.6}$$

The first, second and fifth equations of (8.4) tell us therefore how spherical aberration and circular coma are affected by a change in the position of the diaphragm. In particular, we have that spherical aberration is independent of it; whilst the comatic displacement (13.31) is a linear function of δY_0, the change in the coma coefficients being proportional to $\underline{n}$ and to the spherical aberration coefficient of the $\underline{n}$th order.

PART II. ALGEBRAIC DEVELOPMENTS. IDENTITIES.

§ 14. Plane of Incidence. Invariance of K^*.

(a) So far we have not obtained any expressions for the quantity $\Delta \underline{\Lambda}_j$, which on the one hand exhibit its quasi-invariant nature in explicit form, and which could on the other form the basis for obtaining the series (5.2). We shall therefore study this problem in the next two paragraphs.

If $\underline{x}_p$, $\underline{y}_p$, $\underline{z}_p$ are the co-ordinates of the point of incidence $\underline{P}$, i.e. the intersection point of ray and surface, then the inward unit normal $\underline{n}$ has the components $(1 - x_p/r, - y_p/r, - z_p/r)$, or $(1 - x_p/r, (x_pV - Y)/r, (x_pW - Z)/r$. (14.1)

Let $\underline{e}(\alpha, -\beta, -\gamma)$, $\underline{e}'(\alpha', -\beta', -\gamma')$ be the unit vectors in the directions of the incident and refracted rays respectively. Then since $\underline{e}$, $\underline{e}'$ and $\underline{n}$ are coplanar, and in virtue of the law of refraction there must be an equation of the form

$$\Delta N \underline{e} = \frac{N' - N}{\sigma} \underline{n}, \tag{14.2}$$

where σ is so far undetermined.

Scalar multiplication by $\underline{n}$ of both sides of (14.2) gives

$$\frac{1}{\sigma} = \frac{\Delta N \cos I}{\Delta N} . \tag{14.21}$$

Similarly vectorial multiplication of (14.1) by $\underline{e}$ gives the important result

direction cosines of normal to plane of incidence

$$= N'\underline{e}' \times \underline{e} = \frac{N' - N}{\sigma} \underline{n} \times \underline{e} \quad , \tag{14.3}$$

or in component form, by (14.1) and (4.51),

$$(\beta\gamma' - \beta'\gamma,\ \gamma\alpha' - \gamma'\alpha,\ \alpha\beta' - \alpha'\beta) = \frac{N' - N}{rNN'\sigma}(K^*,\ -C_z^*,\ +C_y^*)\ , \tag{14.4}$$

$$\text{where} \quad K^* = N(\beta Z - \gamma Y)\ , \tag{14.5}$$

$$\therefore \quad K^* = -\frac{\sigma}{\varpi}(\beta\gamma' - \beta'\gamma) \tag{14.51}$$

$$\left.\begin{aligned} C_y^* &= -\frac{\sigma}{\varpi}(\alpha\beta' - \alpha'\beta) \\ C_z^* &= -\frac{\sigma}{\varpi}(\alpha\gamma' - \alpha'\gamma)\ . \end{aligned}\right\} \tag{14.6}$$

Dividing (8.6) by $\alpha\alpha'$ we have at once

$$\Delta \underline{V} = -\frac{\varpi}{\sigma\alpha'}\ \underline{c}\ . \tag{14.7}$$

The co-ordinates of $\underline{P}$ may now be read off from (14.2).

Also, since $\quad \sin^2 I = (\underline{m} \times \underline{e}) \cdot (\underline{m} \times \underline{e})\ ,$

we have by (14.4) the identity

$$(Nr \sin I)^2 = C_y^{*2} + C_z^{*2} + K^{*2}, \tag{14.8}$$

(b) Equation (14.8) shows at once that

$$\Delta K^* = 0\ . \tag{14.9}$$

But since $\quad \underline{H} = l_o\underline{V} - \underline{Y}$ we may write (14.5) in the form

$$K^* = N(\gamma H_y - \beta H_z) \tag{14.91}$$

so that for all j it also obeys the equation

$$K_j^{*\prime} = K_{j+1}^*\ . \tag{14.92}$$

Hence K^* is an optical invariant.

§15. Expression for $\Delta\underline{\Lambda}$, $\Delta\underline{\Lambda}^*$.

(a) According to the definition of $\underline{\Lambda}$ we have by (4.52)

$$\underline{\Lambda} = c_o\underline{V} - u_o\underline{C} \quad . \tag{15.1}$$

$$\therefore \quad \Delta\underline{\Lambda} = c_o\Delta\underline{V} - \underline{C}^*\Delta\frac{u_o}{\alpha} \quad , \text{ by (4.55)}$$

$$= \underline{C}(-\frac{\varpi c_o}{\sigma\alpha'} - \alpha\Delta\frac{u_o}{\alpha}) \text{ , by (14.7) .}$$

But $\qquad -\varpi c = i - i' = u' - u = e$, say. $\qquad$ (15.2)

$$\therefore \Delta\underline{\Lambda} = \underline{C}\left\{\frac{1}{\sigma\alpha'}(u_o' - u_o) - u_o'\frac{\alpha}{\alpha'} + u_o\right\} .$$

Therefore finally, by (4.6),

$$\Delta\underline{\Lambda} = \underline{C}\,\underline{J} \quad , \tag{15.3}$$

where $\qquad \underline{J} = e_o(\frac{1}{\sigma\alpha'} - 1) - u_o'(\frac{\alpha}{\alpha'} - 1)$.

Since σ, α , α' are obviously each of the form $(1 + O(2))$ it follows at once that $\underline{J} = O(2)$.

(b) In the case of semi-canonical variables it may easily be shown that $\Delta\underline{\Lambda}^*$ splits up in the manner indicated by (T4.10), i.e.

$$\Delta\underline{\Lambda}^* = c_o\underline{\Delta} = c_o\Delta(\underline{\beta} + \sin\underline{I}^*) \quad . \tag{15.4}$$

Equation (4.6) here has as its analogue

$$C_z^*\Delta\Lambda_y^* - C_y^*\Delta\Lambda_z^* = K^*\Delta\alpha \quad . \tag{15.5}$$

Also $\qquad \Delta\underline{\Lambda}^* = \Delta(c_o\underline{\beta} - u_o\underline{C}^*)$

$$= c_o\Delta\underline{\beta} - \underline{C}^*\Delta u_o \quad , \text{ by (4.55)}$$

$$= c_o(\Delta\underline{\beta} + \varpi\underline{C}^*) \text{ , by (15.2)}$$

$$= c_o[\Delta\underline{\beta} - \sigma(\alpha\underline{\beta}' - \alpha'\underline{\beta})] \text{, by (14.6) .}$$

91

$$\therefore \quad \underline{\Delta} \quad = \quad \underline{\beta}'(1 - \alpha\sigma) - \underline{\beta}(1 - \alpha'\sigma) \quad , \qquad (15.6)$$

which again shows at a glance that $\underline{\Delta} = O(3)$.

§16. Standard Form of Primary Aberrations.

It was remarked in the introduction (§1.c.) that the present method was of value in connection with the usual theory of primary aberration. When primary terms only are considered all we need to retain of the work of Part I is the paraxial theory and equations (3.6) and (3.71). To obtain the aberrations it is simplest to develop expressions for $\Delta\underline{\Lambda}$ in the manner of (T§8). But once having (15.3) at our disposal we may as well make use of it.

Using small letters to represent the linear terms in (4.2), e.g.

$$\sin I_y \rightarrow i_y \quad = \quad i_p Y_1 + i_q V_1 \quad , \qquad (16.1)$$

we have

$$\alpha \quad = \quad 1 - \tfrac{1}{2}(v^2 + w^2) + \underline{O}(4) \qquad (16.2)$$

and by (14.21)

$$\frac{1}{\sigma} \quad = \quad \frac{N'\sqrt{1 - k^2\sin^2 I} - N\sqrt{1 - \sin^2 I}}{N' - N} \qquad (16.3)$$

$$= \frac{[1 - \frac{1}{2}k^2(i_y{}^2 + i_z{}^2)] - k[1 - \frac{1}{2}(i_y{}^2 + i_z{}^2)]}{1 - k} + \underline{O}(4), \quad \text{by (14.8)}$$

$$= \quad 1 + \tfrac{1}{2}(i_y i_y{}' + i_z i_z{}') + \underline{O}(4) \quad . \qquad (16.31)$$

Substituting these results in (15.3)

$$\underline{J}^{(1)} = \quad \tfrac{1}{2}\left\{e_o\left[i_y i_y{}' + i_z i_z{}' + v'^2 + w'^2\right] - u_o{}'\,\Delta(v^2 + w^2)\right\} \quad . \qquad (16.4)$$

Now if quantities without the suffix $\underline{o}$ refer to the principal ray, then

$$v = u_p(y_1 + \rho y_{o1} \cos\theta) + u_q(y_1 + \rho y_{o1}\cos\theta + h)/l_{o1}$$

$$= u + \rho u_o \cos\theta\ , \qquad (16.5)$$

where ρy_{o1} replaces the ρ of §6.

Also $$\lambda = Nr(i_o u - u_o i) = c_o u - c u_o\ . \qquad (16.51)$$

Hence if we write in this paragraph

$$\left.\begin{aligned} \frac{i}{i_o} &= q \\ q + \rho\cos\theta &= \beta \end{aligned}\right\} \qquad (16.6)$$

we have $$\left.\begin{aligned} v &= \beta u_o + \lambda/c_o \\ \text{and similarly}\quad w &= \rho u_o \sin\theta \\ i_y &= \beta i_o \\ i_z &= \rho i_o \sin\theta\ . \end{aligned}\right\} \qquad (16.61)$$

Substituting (16.61) in (16.4) we find directly

$$\begin{aligned} \underline{J}^{(1)} &= \tfrac{1}{2}\left[(i_o i_o' + u_o'^2)e_o - u_o'\Delta u_o{}^2\right](\beta^2 + \rho^2\sin^2\theta) \\ &\qquad + \left[e_o u_o' - u_o'\Delta u_o\right]\frac{\beta\lambda}{c_o} + \tfrac{1}{2}e_o\frac{\lambda^2}{c_o{}^2} \\ &= \tfrac{1}{2}(i_o - i_o')\left\{(i_o i_o' - u_o u_o')(\beta^2 + \rho^2\sin^2\theta) + \frac{\lambda^2}{c_o{}^2}\right\}\ , \end{aligned} \qquad (16.7)$$

Writing for the contribution to the spherical aberration coefficient

$$S = \tfrac{1}{2}c_o(i_o - i_o')(i_o + u_o)(i_o' - u_o)\ , \qquad (16.71)$$

we have at once, by (15.3), (16.61),

$$\left.\begin{aligned} \Delta\Lambda_y &= (q + \rho\cos\theta)\left\{(q^2 + 2q\rho\cos\theta + \rho^2)S - \tfrac{1}{2}\varpi\lambda^2\right\} \\ \Delta\Lambda_z &= \rho\sin\theta\left\{(q^2 + 2q\rho\cos\theta + \rho^2)S - \tfrac{1}{2}\varpi\lambda^2\right\} \quad . \end{aligned}\right\} \quad (16.8)$$

Hence if we put

$$\left.\begin{aligned} \sigma_1 &= \frac{1}{\mu'}\sum_{j=1}^{k} S_j & \qquad \sigma_4 &= -\frac{1}{2\mu'}\lambda^2\sum_{j=1}^{k}\varpi_j \\ \sigma_2 &= \frac{1}{\mu'}\sum_{j=1}^{k} q_j S_j & \qquad \sigma_5 &= \frac{1}{\mu'}\sum_{j=1}^{k}\left(q_j^3 S_j - \tfrac{1}{2}\varpi_j q_j \lambda^2\right) \\ \sigma_3 &= \frac{1}{\mu'}\sum_{j=1}^{k} q_j^2 S_j & & \end{aligned}\right\} \quad (16.9)$$

we have finally the desired result

$$\left.\begin{aligned} \varepsilon'_y &= \sigma_1\rho^3\cos\theta + \sigma_2\rho^2(2 + \cos 2\theta) + (3\sigma_3 + \sigma_4)\rho\cos\theta + \sigma_5 \\ \varepsilon'_z &= \sigma_1\rho^3\sin\theta + \sigma_2\rho^2\sin 2\theta + (\sigma_3 + \sigma_4)\rho\sin\theta \quad . \end{aligned}\right\} \quad (16.10)$$

As the condition for the absence of the primary aberrations we have at once from (16.10) that σ_i $(i = 1, \ldots, 5)$ must vanish. (v. Whittaker, 1915 (b)). Considering the unified manner in which the σ_i here appear simultaneously it is difficult to understand the precise meaning of the statement (Hardy and Perrin, 1932 (a)) "... the equation for the fifth Seidel aberration has no meaning unless each of the other four is zero." (v. also Whittaker, loc.cit.)

§ 17. Identities.

(a) The simplicity of the present method in its application to practical computation is partly due to the use of a number of aberration coefficients which exceeds the least number of coefficients necessary completely to determine the aberrations. We may therefore expect certain identities to exist between some of the coefficients here employed. Thus, consider first the primary coefficients.

Let the "partial expansion" of $\underline{J}$ at the $\underline{j}$th surface (i.e. the expansion of $\underline{J}$ corresponding to the series (5.2), (5.21), (5.211)) be

$$\underline{J} = \sum_{n=1}^{\infty} \sum_{\mu=0}^{n} \sum_{\nu=0}^{\mu} \hat{\underline{g}}_{\mu\nu}^{(n)} \xi^{n-\mu} \eta^{\mu-\nu} \zeta^{\nu} \quad . \tag{17.1}$$

Then the "partial coefficients" appearing in (5.21) are according to (4.6) given by

$$\left.\begin{aligned} \underline{g}_{\mu\nu}^{(n)} &= c_p \, \hat{\underline{g}}_{\mu\nu}^{(n)} \\ \overline{\underline{g}}_{\mu\nu}^{(n)} &= c_q \, \hat{\underline{g}}_{\mu\nu}^{(n)} \quad . \end{aligned}\right\} \tag{17.12}$$

From (16.4) we have at once

$$\left.\begin{aligned} \hat{\underline{a}} &= \tfrac{1}{2}\{e_o(i_p i_p' + u_p'^2) + u_o'(u_p^2 - u_p'^2)\} \\ \hat{\underline{b}} &= \{e_o(i_p i_q' + u_p' u_q') + u_o'(u_p u_q - u_p' u_q')\} \\ \hat{\underline{c}} &= \tfrac{1}{2}\{e_o(i_q i_q' + u_q'^2) + u_o'(u_q^2 - u_q'^2)\} \quad . \end{aligned}\right\} \tag{17.13}$$

Then, since $\quad c_q i_p i_p' - c_p i_p i_q' = 0 \quad ,$

$$\begin{aligned} 2\underline{a}_q - \underline{b}_q &= e_q\{c_q u_p'^2 - c_p u_p' u_q'\} + u_q'\{c_p(u_p^2 - u_p'^2) - c_q(u_p u_q - u_p' u_q')\} \\ &= N_1\{-e_q u_p' + u_q'(u_p - u_p')\} \quad , \text{by (4.12)} \\ &= N_1(-u_q u_p' + u_q' u_p) = -N_1^2 \varpi \quad , \text{by (4.12)}. \end{aligned} \tag{17.14}$$

Similarly $\quad \overline{b}_p - 2c_p \;=\; \underline{\overline{b}}_p - 2\underline{c}_p \;=\; +N_1{}^2\varpi \;.$ (17.141)

$$\therefore \quad 2\overline{A}_{qj} - B_{qj} \;=\; -\overline{B}_{pj} + 2C_{pj} \;=\; -N_1{}^2 \sum_{\nu=1}^{j-1} \varpi_\nu \qquad (17.15)$$

In the same way we may show that

$$2\overline{A}_{pj} - B_{pj} \;=\; -\overline{B}_{qj} + 2C_{qj} \;=\; 0 \;. \qquad (17.16)$$

Combining (17.15) and (17.16)

$$\left.\begin{aligned} 2\overline{A}_j - B_j \;&=\; -N_1{}^2 u_{01} \sum_{\nu=1}^{j-1} \varpi_\nu \\ \overline{B}_j - 2C_j \;&=\; +N_1{}^2 y_{01} \sum_{\nu=1}^{j-1} \varpi_\nu \;. \end{aligned}\right\} \quad (17.17)$$

Hence in particular we have amongst the six primary coefficients for a given position of the object the single identity

$$\Omega^{(1)}_{\infty} = (2\overline{A}_j - B_j)y_{01} + (\overline{B}_j - 2C_j)u_{01} \;\equiv\; 0 \qquad (17.18)$$

for every j; and in particular we have of course

$$(2\overline{A}_k{}' - B_k{}')y_{01} + (\overline{B}_k{}' - 2C_k{}')u_{01} \;=\; 0 \;. \qquad (17.181)$$

By the same method we find without difficulty

$$\left.\begin{aligned} \overline{\underline{a}}_p - \underline{a}_q \;&=\; \tfrac{1}{2} N_1 \Delta u_p{}^2 \\ \overline{\underline{b}}_p - \underline{b}_q \;&\; N_1 \Delta u_p u_q \\ \overline{\underline{c}}_p - \underline{c}_q \;&=\; \tfrac{1}{2} N_1 \Delta u_q{}^2 \;. \end{aligned}\right\} \quad (17.19)$$

Summing these from 1 to $j - 1$ we obtain

$$\left.\begin{aligned} \overline{A}_{pj} - A_{qj} &= \tfrac{1}{2} N_1 u_{pj}^{\;2} \\ \overline{B}_{pj} - B_{qj} &= N_1 u_{pj} u_{qj} \\ \overline{C}_{pj} - C_{qj} &= \tfrac{1}{2} N_1 (u_{qj}^{\;2} - 1) \quad . \end{aligned}\right\} \qquad (17.191)$$

(17.15), (17.16) and (17.191) constitute the required <u>six</u> identities between the general <u>twelve</u> primary coefficients. (17.191) also shows that if the primary aberrations are to vanish identically for arbitrary positions of the object the system must be telescopic with unit reduced magnification, ($u_{pk}{}' = 0$, $u_{qk}{}' = 1$). (see also next section).

(b) A certain set of identities of the type (17.191) may be derived up to all orders by making use of the invariance of K^*, (§14.b). By (14.5), (4.3) and using (4.11)

$$K_j = K_j^*/\alpha_j = N_j \begin{vmatrix} V_j & Y_j \\ W_j & Z_j \end{vmatrix} = N_1 \begin{vmatrix} V_1 + \delta_{vj} \;, & Y_1 + \delta_{yj} \\ W_1 + \delta_{wj} \;, & Z_1 + \delta_{zj} \end{vmatrix} \; . \qquad (17.2)$$

Expressing the increments in terms of the series (4.4), it will be seen that the right-hand side of (17.2) contains K_1 as a factor, <u>viz.</u>

$$K_j = K_1 + \frac{K_1}{N_1}\left\{(\overline{A}_{pj} - A_{qj})\,\xi + (\overline{B}_{pj} - B_{qj})\,\eta + (\overline{C}_{pj} - C_{qj})\,\varsigma + \underline{O}(4)\right\} \qquad (17.21)$$

Dividing throughout by K_1^* $(=K_j^*)$ we have

$$\frac{\alpha_1}{\alpha_j} - 1 = \frac{1}{N_1}\left\{ \;\cdots\; \right\} . \qquad (17.22)$$

Expanding $\dfrac{\alpha_1}{\alpha_j} = (1 + V_j^{\,2} + W_j^{\,2})^{\frac{1}{2}}\,(1 + \varsigma)^{-\frac{1}{2}}$ as a series in ξ, η, ς, the terms of the second degree in (17.21) give us again the result (17.191). For the higher orders the algebraic work is straightforward but somewhat tedious. We shall merely quote the explicit results obtained by considering the terms of the fourth degree, which express certain combinations of the secondary coefficients in terms of those of lower order alone. For this purpose we write

$$\left.\begin{array}{l} \dfrac{1}{N_1}\begin{vmatrix} A_p & A_q \\ \bar{A}_p & \bar{A}_q \end{vmatrix} = [A\bar{A}] = -[\bar{A}A] \;; \\ \text{and analogously for other pairs of coefficients; also write} \\ \begin{vmatrix} A_p & A_q \\ u_p & u_q \end{vmatrix} = [A] \;; \;\ldots\ldots \end{array}\right\} \quad (17.23)$$

Then collectively denoting by ω_1, ω_2, ... the terms containing the coefficients of lower order only we have

$$\left.\begin{array}{lll} \bar{S}_{1p} - S_{1q} &= [\bar{A}A] + u_p[A] - \tfrac{1}{8}N_1 u_p^4 &= \omega_1 \\ \bar{S}_{2p} - S_{2q} &= [\bar{A}B] + [\bar{B}A] + u_q[A] + u_p[\bar{A}] + u_p[B] - \tfrac{1}{2}N_1 u_p^3 u_q &= \omega_2 \\ \bar{S}_{3p} - S_{3q} &= [\bar{A}C] + [\bar{C}A] + u_q[\bar{A}] + u_p[C] - \tfrac{1}{4}N_1 u_p^2(u_q^2 + 1) &= \omega_3 \\ \bar{S}_{4p} - S_{4q} &= [\bar{B}B] + u_q[B] + u_p[\bar{B}] - \tfrac{1}{2}N_1 u_p^2 u_q^2 &= \omega_4 \\ \bar{S}_{5p} - S_{5q} &= [\bar{B}C] + [\bar{C}B] + u_q[\bar{B}] + u_q[C] + u_p[\bar{C}] - \tfrac{1}{2}N_1 u_p u_q(u_q^2+1) &= \omega_5 \\ \bar{S}_{6p} - S_{6q} &= [\bar{C}C] + u_q[\bar{C}] - \tfrac{1}{8}N_1(u_q^2 + 3)(u_q^2 - 1) &= \omega_6 \end{array}\right\} \quad (17.24)$$

These identities will of course apply before or after <u>any</u> surface. More generally, it will be seen that the expressions

$$\bar{G}^{(n)}_{\mu\nu|p} - G^{(n)}_{\mu\nu|q} \qquad \begin{pmatrix} \mu = 0, 1, \ldots, n \\ \nu = 0, 1, \ldots, \mu \end{pmatrix} \qquad (17.25)$$

must express themselves in terms of the aberration coefficients of order less than <u>n</u> alone; and the number of such identities will be

$$\tfrac{1}{2}(n+1)(n+2) \quad , \quad (n = 1, 2, \ldots) \quad . \tag{17.26}$$

An interesting result follows at once from (17.2) by noticing that if the optical system be perfect the final increments must vanish identically; so that

$$K_k' = K_1 \quad . \tag{17.27}$$

But $$K_k^* = K_1^* \text{ , always.} \tag{17.271}$$

Hence in this case $\alpha_1 \equiv \alpha_k'$. (17.28)

Since (17.28) also implies that $u_{01} \equiv u_{0k}'$ we have the general result that <u>a perfect optical system must be telescopic</u> with reduced magnification unity. This is a theorem concerned with the special type of optical system here considered; a proof of the most general case may be found in Whittaker, (<u>Whittaker</u>, 1915 (c)).

(c) To determine the higher order identities of the type (17.15), (17.16) by a direct algebraic procedure, though no doubt possible, would require work so tedious as hardly to be worth while when it is remembered that these identities play no fundamental part in our theory. Indeed even their practical use is confined to a few arithmetical checks (v.§26) of doubtful value. However, for the sake of completeness we shall briefly discuss these identities in this section by means of a slightly different method. By (15.1), (14.6)

$$\underline{H} = \frac{(l_o - r)\underline{\beta}}{\alpha} + \frac{\sigma}{N\varpi}(\underline{\beta}' - \frac{\alpha'}{\alpha}\underline{\beta}) \quad . \tag{17.3}$$

$$\text{Now } \cos(I - I') = \underline{e} \cdot \underline{e}' = \alpha\alpha' + \beta\beta' + \gamma\gamma' \quad . \tag{17.31}$$

Therefore, by (2.5)

$$\sigma = \frac{N' - N}{\{N'^2 + N^2 - 2NN'(\alpha\alpha' + \beta\beta' + \gamma\gamma')\}^{\frac{1}{2}}} \quad . \tag{17.32}$$

Regarding now β , β', γ, γ' as independent variables it is easily verified by direct differentiation of (17.3) that the relations of the kind

$$\frac{\partial H_y'}{\partial \gamma'} = \frac{\partial H_z'}{\partial \beta'} \tag{17.33}$$

necessary to ensure that the expression

$$N'(H_y' d\beta' + H_z' d\gamma') - N(H_y d\beta + H_z d\gamma) \quad , \tag{17.34}$$

shall be a total differential, are identically satisfied. ((17.34) plays a central role in Hamilton's method). Since this is true at every surface it follows, after addition of a total differential, that if the object point is not varied

$$\beta_j d\Lambda_{yj} + \gamma_j d\Lambda_{zj} = \beta_j dD_{yj} + \gamma_j dD_{zj} = \text{total differential} \; , \tag{17.35}$$

where we now use canonical co-ordinates for specifying rays. Hence, omitting the suffix j,

$$\frac{\partial\beta}{\partial Z_1}\frac{\partial D_y}{\partial Y_1} - \frac{\partial\beta}{\partial Y_1}\frac{\partial D_y}{\partial Z_1} = \frac{\partial\gamma}{\partial Y_1}\frac{\partial D_z}{\partial Z_1} - \frac{\partial\gamma}{\partial Z_1}\frac{\partial D_z}{\partial Y_1} \quad , \tag{17.36}$$

where during the differentiations we imagine V_1 to be written as

$$\underline{V}_1 = (\underline{Y}_1 + \underline{h})/l_{o1} \quad . \tag{17.361}$$

Considering explicitly only terms up to the fourth degree, (17.36) becomes

$$\frac{\partial V}{\partial Y_1}\frac{\partial D_y}{\partial Z_1} - \frac{\partial W}{\partial Z_1}\frac{\partial D_z}{\partial Y_1} = \frac{\partial V}{\partial Z_1}\frac{\partial D_y}{\partial Y_1} - \frac{\partial W}{\partial Y_1}\frac{\partial D_z}{\partial Z_1}$$
$$+ \frac{u_o}{y_{o1}}\left\{V^2\frac{\partial D_z}{\partial Y_1} - W^2\frac{\partial D_y}{\partial Z_1} + VW\left(\frac{\partial D_z}{\partial Z_1} - \frac{\partial D_y}{\partial Y_1}\right)\right\} + \underline{O}(6) \quad . \tag{17.37}$$

Notice that the right-hand side does not contain terms of the second degree. Substituting the series for $\underline{V}$ and $\underline{D}$ in (17.37) comparison of the coefficients of the second degree reproduces the identity (17.18).

It is scarcely necessary to give the somewhat tedious though straightforward derivation of the second order identities in detail. There are three identities of the type (17.18), of which the first after simplification by means of (17.18) is

$$\Omega^{(2)}_{oo} = (4\bar{S}_1 - S_2)y_{o1} + (\bar{S}_2 - 2S_3)u_{o1} \equiv 2u_p u_q A y_{o1} + (-u_p^2 y_{o1} + u_p u_q u_{o1})B$$
$$- 2u_p^2 C u_{o1} + \left\{2[\bar{A}A] + 3[AB]\right\} y_{o1}^2 + \left\{[\bar{B}A] + 6[AC] + [\bar{A}B]\right\} y_{o1} u_{o1}$$
$$+ \left\{[\bar{B}A] + 2[\bar{A}C] + 2[BC]\right\} u_{o1}^2 \quad . \tag{17.4}$$

However, y_{o1}, u_{o1} may now be looked upon as independent variable parameters, so that (17.4) splits up into three component equations. Then the coefficients of y_{o1}^2, u_{o1}^2 give respectively, after simplification by means of the first order identities,

(i) $$4\bar{S}_{1p} - S_{2p} = 2[\bar{A}A] + 3[AB] + 2u_p[A] - N_1 u_p^4 = \omega_7 \tag{17.41}$$

(ii) $$\bar{S}_{2q} - 2S_{3q} = [\bar{B}\bar{A}] + 2[\bar{A}C] + 2[BC] + u_p[\bar{B}] - N_1 u_p^2 u_q^2 = \omega_8 \tag{17.42}$$

The coefficient of $y_{o1}u_{o1}$ similarly gives

$$4\bar{S}_{1q} - S_{2q} + \bar{S}_{2p} - 2S_{3p} = [\bar{B}A] + 6[AC] + [\bar{A}B] + 2u_p[\bar{A}]$$

$$+ u_p[B] - 2N_1 u_p^3 u_q \quad .$$

This may be further simplified by means of (17.24); thus subtracting ω_2 from both sides we find

$$\text{(iii)}\ 4\bar{S}_{1q} - 2S_{3p} = 6[AC] - u_q[A] + u_p[\bar{A}] - \tfrac{3}{2}N_1 u_p^3 u_q = \omega_9 \tag{17.43}$$

The second identity may in a similar way be shown to be

$$\Omega_{10}^{(2)} = (2\bar{S}_2 - 2S_4)y_{01} + (2\bar{S}_4 - 2S_5)u_{01} \equiv 2u_q^2 A y_{01} + u_q^2 B u_{01}$$

$$- u_p^2 \bar{B} y_{01} - 2u_p^2 \bar{C} u_{01} + \left\{2[A\bar{B}] + 4[\bar{A}B]\right\} y_{01}^2 + \left\{2[\bar{A}\bar{B}] + 4[\bar{A}C]\right.$$

$$\left. + 2[\bar{B}B] + 4[A\bar{C}] + 2[BC]\right\} y_{01} u_{01} + \left\{4[\bar{B}C] + 2[B\bar{C}]\right\} u_{01}^2 \tag{17.5}$$

which, again, splits up into

$$\text{(i)}\ 2\bar{S}_{2p} - 2S_{4p} = 2[A\bar{B}] + 4[\bar{A}B] + 2u_q[A] + u_p[B] - 2N_1 u_p^3 u_q = \omega_{10} \tag{17.51}$$

$$\text{(ii)}\ 2\bar{S}_{4q} - 2S_{5q} = 4[\bar{B}C] + 2[B\bar{C}] + u_q[\bar{B}] + 2u_p[\bar{C}] - N_1 u_p u_q(2u_q^2 - 1) = \omega_{11} \tag{17.52}$$

$$\text{(iii)}\ 2\bar{S}_{2q} - 2S_{5p} = 2[\bar{A}\bar{B}] + 4[\bar{A}C] + 4[A\bar{C}] + 2[BC] - N_1 u_p^2(3u_q^2 - 1) = \omega_{12} \tag{17.53}$$

Finally the third identity

$$\Omega_{11}^{(2)} = (2\bar{S}_3 - S_5)y_{01} + (\bar{S}_5 - 4S_6)u_{01} \equiv 2u_q^2 \bar{A} y_{01} + (-u_p u_q y_{01} + u_q^2 u_{01})\bar{B}$$

$$- 2u_p u_q \bar{C} u_{01} + \left\{2[\bar{A}\bar{B}] + 2[\bar{A}C] + [CB]\right\} y_{01}^2 + \left\{[\bar{C}B + 6[\bar{A}\bar{C}] + [\bar{B}C]\right\} y_{01} u_{01}$$

$$+ \left\{3[\bar{B}\bar{C}] + 2[\bar{C}C]\right\} u_{01}^2 \tag{17.6}$$

which splits up into

$$\text{(i)}\ 2\bar{S}_{3p} - S_{5p} = 2[\bar{A}\bar{B}] + 2[\bar{A}C] + [CB] + u_q[\bar{B}] - N_1 u_p^2 u_q^2 = \omega_{13} \tag{17.61}$$

(ii) $\bar{S}_{5q}-4S_{6q} = 3[\bar{B}\bar{C}]+2[\bar{C}C]+2u_q[\bar{C}] - N_1 u_q^2(u_q^2-1) = \omega_{14}$ (17.62)

(iii) $2\bar{S}_{3q}-4S_{6p} = 6[\bar{A}\bar{C}]+u_q[C]-u_p[\bar{C}]-\frac{3}{2}N_1u_pu_q(u_q^2-1) = \omega_{15}$ (17.63)

More generally it will be seen that there must be

$$\tfrac{1}{2}n(n+1) \qquad (n=1, 2, \ldots) \tag{17.7}$$

such identities between the coefficients of order $\underline{n}$, and these are of the form

$$\Omega^{(n)}_{\mu\nu} = \left(2(n-\mu)\bar{G}^{(n)}_{\mu\nu} - (\mu-\nu+1)G^{(n)}_{\mu+1,\nu}\right)y_{01} + \left((\mu-\nu+1)\bar{G}^{(n)}_{\mu+1,\nu} - 2(\nu+1)G^{(n)}_{\mu+1,\nu+1}\right)u_{01}$$

$\equiv$ terms involving only coefficients of order less than $\underline{n}$. (17.71)

$$(\mu = 0, 1, \ldots, n-1; \quad \nu = 0, 1, \ldots, \mu)$$

Each of these then splits up, as above, into three component identities.

(d) From the 15 second order identities above it will be seen that we can express the difference $\bar{S}_{2q} - S_{5p}$ in two ways, i.e. either as $\frac{1}{2}\omega_{12}$; or as $\omega_{13}+\omega_8-2\omega_3$, so that we must have

$$4\omega_3 - 2\omega_8 + \omega_{12} - 2\omega_{13} \equiv 0 \quad , \tag{17.8}$$

for otherwise a further identity between the primary coefficients would be implied. It is easily verified that (11.8) is satisfied. Similarly by a slight modification of the method of section (c) it may be shown that $\frac{1}{2}n(n-1)$ of the $\underline{n}$th order identities are redundant.

Hence the total number of identities is by (17.26), (17.7)

$$\tfrac{1}{2}(n+1)(m+2) + \tfrac{3}{2}m(m+1) - \tfrac{1}{2}m(m-1) \quad = \quad \tfrac{1}{2}(3m+1)(m+2) \quad . \qquad (17.81)$$

But, by (3.94) there are $2(n+1)(n+2)$ $\underline{n}$th order coefficients

$\therefore$ The number of independent coefficients of order $\underline{n}$

$$\tfrac{1}{2}(m + 2)(m + 3) \quad . \qquad (17.82)$$

If the position of the object is fixed (v. Steward, 1928 (b)) the number of coefficients is by (3.94) and (17.7)

$$= (m+1)(n+2) - \tfrac{1}{2}n(n+1) \quad = \quad \tfrac{1}{2}(n+1)(n+4) \, . \qquad (17.83)$$

The difference between (17.81) and (17.83) is unity. Hence the number of independent aberration coefficients of order $\underline{n}$ for arbitrary positions of the object exceeds by one the number of coefficients for a fixed position of the object.

(e) When general linear variables are being used certain identities will exist between the coefficients (12.72). But on account of the greater complexity of (12.41) which causes all the coefficients $F^{(n)}_{\mu\nu,j|\alpha}$, ..., $(j = 1, 2, \ldots, k)$ to appear at every surface the derivation of the identities in the most general case becomes very tedious. We therefore content ourselves with a simple but useful special case, viz.

$$\sigma = \tau = \bar{\sigma} = \bar{\tau} = 0 \quad .$$

Then $\underline{\delta}_{sj}$, $\underline{\delta}_{tj}$ are given by (12.52), with $g = \begin{vmatrix} \pi & \rho \\ \bar{\pi} & \bar{\rho} \end{vmatrix}$.

Hence, remembering that now u_{a1} is not necessarily zero we find at once by comparison with (17.191) the first order identities

$$\left.\begin{aligned}\overline{F}^{(1)}_{00,j|a} - F^{(1)}_{00,j|b} &= \tfrac{1}{2}N_1(u_{aj}{}^2 - u_{a1}{}^2)/g \\ \overline{F}^{(1)}_{10,j|a} - F^{(1)}_{10,j|b} &= N_1(u_{aj}u_{bj} - u_{a1}u_{b1})/g \\ \overline{F}^{(1)}_{11,j|a} - F^{(1)}_{11,j|b} &= \tfrac{1}{2}N_1(u_{bj}{}^2 - u_{b1}{}^2)/g \quad .\end{aligned}\right\} \quad (17.9)$$

In place of (17.18) we have now

$$\widetilde{\Omega}^{(1)}_{00} = (2\overline{F}^{(1)}_{00} - F^{(1)}_{10})s_0 + (\overline{F}^{(1)}_{10} - 2F^{(1)}_{11})t_0 \equiv 0 \ . \quad (17.91)$$

(Actually this holds also in the general case; and it may be shown that in a change to general linear co-ordinates $\Omega^{(1)}_{00}$ transforms into $g\,\widetilde{\Omega}^{(1)}_{00}$). In the same way, analogous expressions may be written down for $\widetilde{\Omega}^{(2)}_{00}$ $\widetilde{\Omega}^{(2)}_{10}$, $\widetilde{\Omega}^{(2)}_{11}$ by inspection of (17.4 - 6) if N_1 be replaced by N_1/g and all other symbols by the present generalised symbols; and similarly for higher orders.

As a further specialisation we may consider the case (12.5); we then have the additional simplification

$$t_0 \equiv 0 \ . \quad (17.92)$$

Consequently we obtain the first of the second order identities in the form

$$\widetilde{\Omega}^{(2)}_{00}/s_0 = (4\overline{F}^{(2)}_{00} - F^{(2)}_{10}) \equiv 2u_a u_b F^{(1)}_{00} - u_a^2 F^{(1)}_{10} + \{2g[\overline{F}^{(1)}_{00} F^{(1)}_{00}] + 3g[F^{(1)}_{00} F^{(1)}_{10}]\} , \quad (17.93)$$

an equation which has already been used in §13.b.

In the case of semi-canonical co-ordinates analogous considerations will yield identities not essentially different from those derived above.

PART III. EXPLICIT SERIES EXPANSIONS. SAMPLE COMPUTATIONS.

§18. On Alternative Methods of Calculation.

In order to carry out the iteration explained in §5. we need to determine the explicit expressions for the $\underline{\Gamma}_{n,j}$ of equation (5.2). This will be done below. But it is important to remember that $\underline{\Gamma}_{n,j}$ must there be expressed entirely in terms of the canonical variables before the surface in question; i.e. $\underline{Y}_j'$, ..., may not occur. This result leads to a certain complexity in the resulting expansions mostly on account of the involved dependence of α_j' on $\underline{Y}_j$, $\underline{V}_j$ (v. §19d).

Accordingly we may also proceed in a somewhat different manner (though the same in principle as that above) which may conceivably be applied more simply in practice; such as, perhaps, in the computation of quaternary or higher spherical aberration. In this method we expand $\Delta\underline{\Lambda}_j$ in the form

$$\Delta\underline{\Lambda}_j = \sum_{n=1}^{\infty} \underline{\bar{\Gamma}}_{n,j}(Y_j, Y_j', Z_j, Z_j', V_j, V_j', W_j, W_j') \qquad , \tag{18.1}$$

i.e. we allow in $\underline{\bar{\Gamma}}_{n,j}$ the canonical variables before and after refraction to occur, in such a way as to reduce the number of terms in it to a minimum. We then obtain an expansion equivalent to (5.21) by setting for any canonical variable $\underline{A}_j$ before refraction

$$\underline{A}_j = \alpha_{pj}\ \underline{Y}_1 + \alpha_{qj}\ \underline{V}_1, \tag{18.2}$$

and for the corresponding variable after refraction

$$\underline{A}_j' = \alpha_{pj}' \, \overline{\underline{Y}}_1 + \alpha_{qj}' \, \overline{\underline{V}}_1 \quad . \tag{18.21}$$

When inserting increments subsequently we must replace $\underline{Y}_1$ by $\underline{Y}_1 + \underline{\delta}_{yj}$, whilst $\overline{\underline{Y}}_1$ must be replaced by $\overline{\underline{Y}}_1 + \underline{\delta}'_{yj}$; and similarly for $\underline{V}_1, \overline{\underline{V}}_1$. When the required order has been reached the bars may finally be dropped. Notice that if we require the aberrations correct to the nth order the difference between barred and unbarred co-ordinates may be ignored in the expansion of $\overline{\underline{\Gamma}}_{n,j}$.

This alternative method will also be briefly considered below (§24).

§19. First Method in General. Canonical Co-ordinates.

(a) In order to determine the $\underline{\Gamma}_{n,j}$ explicitly it is necessary, by (4.6), to expand $\underline{J}_j$ explicitly in terms of the canonical variables at (i.e. before) the jth surface. We begin by rewriting $\underline{J}$ in the form

$$\underline{J} = e_o \frac{\alpha}{\alpha'}\left(\frac{1}{\alpha\sigma} - 1\right) - u_o\left(\frac{\alpha}{\alpha'} - 1\right) \quad . \tag{19.1}$$

If we put

$$\left.\begin{aligned} \frac{1}{\alpha\sigma} - 1 &= \underline{S} \\ \frac{\alpha}{\alpha'} - 1 &= \underline{T} \quad , \end{aligned}\right\} \tag{19.11}$$

this becomes

$$\underline{J} = e_o(1 + \underline{T})\underline{S} - u_o\underline{T} \quad . \tag{19.12}$$

If a superscript in brackets (n) indicates a homogeneous polynomial of degree 2n, then

$$\left.\begin{aligned} \underline{S} &= \sum_{n=1}^{\infty} S^{(n)} \\ \underline{T} &= \sum_{n=1}^{\infty} T^{(n)} \\ \underline{J} &= \sum_{n=1}^{\infty} J^{(n)} \end{aligned}\right\} \quad (19.2)$$

so that

$$\left.\begin{aligned} J^{(1)} &= e_0 S^{(1)} - u_0 T^{(1)} \\ J^{(n)} &= e_0 S^{(n)} - u_0 T^{(n)} + e_0 \sum_{\nu=1}^{n-1} S^{(\nu)} T^{(n-\nu)}, \quad (n>1) \end{aligned}\right\} \quad (19.21)$$

and finally

$$\underline{\Gamma_{n,j}} = \underline{c}\, J_j^{(n)} . \qquad (19.22)$$

(b) We shall find it most convenient below to introduce the following notation:

$$\left.\begin{aligned} E &= k(\sin^2 I_y + \sin^2 I_z) \\ F &= \frac{(1-k)^2}{k} E = E/\kappa, \text{ say} \\ G &= (1-k)(V \sin I_y + W \sin I_z) \\ H &= V^2 + W^2 \end{aligned}\right\} \quad (19.3)$$

and the following combinations of these

$$\left.\begin{aligned} H - \frac{k^2 F}{(1-k)^2} &= \underline{F} \\ H - \frac{F}{(1-k)^2} &= \underline{G} \\ \frac{G^2 - FH}{(1-k)^2} &= \underline{K}^2 = -(K/Nr)^2 . \end{aligned}\right\} \quad (19.31)$$

The numerous symbols introduced amply justify themselves in preventing the appearance of an unmanageable jungle of different terms.

In (19.3) $\sin \underline{I}_y, \underline{V}, \ldots$ will eventually be replaced by their paraxial equivalents (i.e. in passing from (5.1) to (5.2). To indicate this clearly we then write $\sin \underline{I}_y = \underline{i}_y, \underline{V} = \underline{v}, \ldots$. Then, for example, G becomes

$$G = (1-k)(vi_y + wi_z) = ve_y + we_z$$

which we shall then write simply as

$$\left.\begin{aligned} G &= ue \quad (\rightarrow u_p\, e_p \xi + (u_p e_q + u_q e_p)\eta + u_q e_q \varsigma) \\ \text{and similarly} \quad E &= ii' \\ F &= e^2 \\ H &= u^2 \end{aligned}\right\} \quad (19.32)$$

Care must be exercised in the use of these symbols. Thus $FH - G^2$ is not zero but

$$(u_p e_q - u_q e_p)^2(\xi\varsigma - \eta^2) = N_1^2 \varpi^2(\xi\varsigma - \eta^2)$$

(c) By (14.8),

$$\sin^2 I = \alpha^2(\sin^2 I_y + \sin^2 I_z + K^2/N^2 r^2)$$

$$\therefore \sin^2 I = \frac{F/(1-k)^2 - \underline{K}^2}{1 + H}$$

Hence by (16.3)

$$\frac{1-k}{\alpha\sigma} = (1-k)(1+\underline{S}) = \sqrt{1 + \underline{F} + k^2\underline{K}^2} - k\sqrt{1 + \underline{G} + \underline{K}^2} \quad . \qquad (19.4)$$

Since $\underline{F}$ and $\underline{G}$ are of the second degree and $\underline{K}^2$ of the fourth degree, (19.4) lends itself readily to expansion in such a way that all terms of the same degree are grouped together (v. also § 24a).

(d) To obtain the expansion of $\underline{T}$ we make use of (14.7). Thus

$$\frac{1}{\alpha'^2} = 1 + V'^2 + W'^2 = 1 + (V - \frac{\varpi C_y}{\alpha'})^2 + (W - \frac{\varpi C_z}{\alpha'})^2 \quad . \tag{19.5}$$

$$\therefore \quad \frac{1}{\alpha'^2} = \frac{1}{\alpha^2} + 2G(1+\underline{S})(\frac{\alpha}{\alpha'}) + F^2(1+\underline{S})^2(\frac{\alpha}{\alpha'})^2 \quad . \tag{19.51}$$

Hence we obtain for $\underline{T}$ the following quadratic equation

$$\left(\frac{F(1+\underline{S})^2}{1 + H} - 1\right)(\underline{T}+1)^2 + \frac{2G(1+\underline{S})}{1 + H} \cdot (\underline{T}+1) + 1 = 0 \quad ; \tag{19.6}$$

The solution of which is

$$\underline{T} = \frac{G(1+\underline{S}) + \quad (1+H)^2 - (1+\underline{S})^2(F + FH - G^2)}{1 + H - F(1+\underline{S})} - 1 \quad . \tag{19.7}$$

The terms $FH - G^2$ in (19.7) may also be written as ϖ^2K^2 whenever convenient. Using (19.4), (19.7) may be expanded directly to give $T^{(n)}$.

§ 20. First Method ctd. Lower Orders in Canonical Co-ordinates.

(a) In this paragraph we employ the results of § 19 in order chiefly to consider terms of the first and second order. It is well worth while to attempt to group the many terms obtained in the expansion of (19.4) and (19.7) in certain sets so as to reduce the number of separate terms to be calculated to a minimum, partly by arranging those of the second order in a way such that as many as possible of the first order terms are repeated in them. This has been done below. But an investigation of the tertiary and higher terms along the same lines would probably amply justify the cumbersome labour involved in finding once for all suitable combinations of them.

(b) Considering throughout only terms up to the sixth degree (second order) we see at once on expanding (19.4) that

$$\left.\begin{aligned} (1-k)S^{(1)} &= \tfrac{1}{2}(\underline{F} - k\underline{G}) \\ (1-k)S^{(2)} &= -\tfrac{1}{8}\left\{(\underline{F}^2 - k\underline{G}^2) + 4k(1-k)\underline{K}^2\right\} \;. \end{aligned}\right\} \quad (20.1)$$

Applying (19.31), and remembering that $E = kF$, these are easily re-expressed in the form

$$\left.\begin{aligned} S^{(1)} &= \tfrac{1}{2}(E + H) \\ S^{(2)} &= \tfrac{1}{8}\left\{(E + H)(3E - H) + k(F + 2G)(F - 2G)\right\}. \end{aligned}\right\} \quad (20.11)$$

To the required order (19.7) now becomes

$$\underline{T} + 1 = \frac{G + GS^{(1)} + \{1 + (2H-F) + (H^2 + G^2 - FH - 2FS^{(1)})\}^{1/2}}{1 + (H-F) - 2FS^{(1)}} \,. \quad (20.2)$$

The square root may be expanded by means of the binomial theorem. Collecting terms of like degree we find

$$\left.\begin{aligned} T^{(1)} &= \tfrac{1}{2}(F + 2G) \\ T^{(2)} &= \tfrac{1}{8}(F + 2G)(3F + 2G - 4H) + \tfrac{1}{2}(F + G)(E + H) \;. \end{aligned}\right\} \quad (20.21)$$

In order further to simplify the computations we re-express E, F, G, H in terms of four other quantities $\gamma_1, \gamma_2, \gamma_3, \gamma_4$; viz.

$$\left.\begin{aligned} E &= \tfrac{1}{2}(\gamma_1 + 3\gamma_2 + 2\gamma_3) \\ F &= -(\gamma_2 + 2\gamma_4) \\ G &= -\tfrac{1}{2}(\gamma_2 - 2\gamma_4) \\ H &= \tfrac{1}{2}(3\gamma_1 + \gamma_2 - 2\gamma_3) \;. \end{aligned}\right\} \quad (20.3)$$

10'

Then we have

$$\left.\begin{aligned}
\tfrac{1}{2}(E + H) &= \gamma_1 + \gamma_2 \\
\tfrac{1}{4}(3E - H) &= \gamma_2 + \gamma_3 \\
\tfrac{1}{2}(F + 2G) &= -\gamma_2 \\
\tfrac{1}{4}(F - 2G) &= -\gamma_4 \\
3F + 2G - 4H &= -6(\gamma_1 + \gamma_2) + 4(\gamma_3 - \gamma_4) \\
F + G &= -(\tfrac{3}{2}\gamma_2 + \gamma_4) .
\end{aligned}\right\} \quad (20.31)$$

We then obtain (20.11) and (20.21) in the simple form

$$\left.\begin{aligned}
S^{(1)} &= \gamma_1 + \gamma_2 \\
S^{(2)} &= (\gamma_1 + \gamma_2)(\gamma_2 + \gamma_3) + k\gamma_2\gamma_4 \\
T^{(1)} &= -\gamma_2 \\
T^{(2)} &= -(\gamma_1\gamma_4 + \gamma_2\gamma_3) ;
\end{aligned}\right\} \quad (20.4)$$

$$\text{then} \quad J^{(1)} = e_o S^{(1)} - u_o T^{(1)} = e_o(\gamma_1 + \gamma_2) + u_o\gamma_2$$

$$= e_o\gamma_1 + u_o'\gamma_2 ;$$

and

$$J^{(2)} = e_o(S^{(2)} + S^{(1)}T^{(1)}) - u_o T^{(2)}$$

$$= e_o\{(\gamma_1 + \gamma_2)(\gamma_2 + \gamma_3) + k\gamma_2\gamma_4 - (\gamma_1 + \gamma_2)\gamma_2\} + u_o(\gamma_1\gamma_4 + \gamma_2\gamma_3)$$

$$= \gamma_3\{e_o(\gamma_1 + \gamma_2) + u_o\gamma_2\} + \gamma_4\{u_o\gamma_1 + ke_o\gamma_2\} .$$

Hence finally

$$J^{(1)} = e_o\gamma_1 + u_o'\gamma_2 \quad (20.41)$$

$$J^{(2)} = \gamma_3 J^{(1)} + \gamma_4(u_o\gamma_1 + ke_o\gamma_2) . \quad (20.42)$$

The γ's will be found from (19.32) and (20.3) to have the simple significance

$$\left.\begin{aligned} \gamma_1 &= \tfrac{1}{2}(ii' + u'^2) \\ \gamma_2 &= \tfrac{1}{2}(u^2 - u'^2) \\ \gamma_3 &= \tfrac{1}{4}(3ii' + 2u'^2 - 3u^2) = \tfrac{3}{2}\gamma_1 + \gamma_4 - uu' \\ \gamma_4 &= -\tfrac{1}{4}(u'^2 - 4uu' + 3u^2) \quad . \end{aligned}\right\} \quad (20.43)$$

Here the γ's are to be thought of as expressed in terms of ξ, η, ζ ; e.g.

$$\gamma_1 = \tfrac{1}{2}(i_p i_p' + u_p'^2)\xi + (i_p i_q' + u_p' u_q')\eta + \tfrac{1}{2}(i_q i_q' + u_q'^2)\zeta; \quad (20.44)$$

whilst finally in (19.22)

$$\underline{c} = c_p \underline{V_1} + c_q \underline{V_1} \quad .$$

(c) The plane surface is most easily dealt with ab initio.

Write $$\Delta\underline{\Lambda} = \Delta N(y_o \underline{V} - u_o \underline{Y}) \quad . \quad (20.5)$$

But now surface and tangent plane coincide.

$$\therefore \quad \Delta\underline{Y} = 0$$

Also $$\Delta N u_o = -\Delta N i_o = 0$$

$$\therefore \quad \Delta\underline{\Lambda} = y_o \Delta N \underline{V} \quad (20.51)$$

Since $$\sin \underline{I} = -\underline{V} \quad (20.52)$$

$$\therefore \quad \underline{V}' = k\frac{\alpha}{\alpha'}\underline{V} \quad .$$

$$\therefore \quad \Delta\underline{\Lambda} = y_o(N'k\frac{\alpha}{\alpha'}\underline{V} - N\underline{V})$$

$$= Ny_o\underline{V}\left(\frac{\alpha}{\alpha'} - 1\right) \quad . \tag{20.53}$$

Now
$$\frac{1}{\alpha'^2} = 1 + V'^2 + W'^2 = 1 + \frac{k^2\alpha^2}{\alpha'^2}(V^2 + W^2)$$

$$= 1 + \frac{k^2\alpha^2}{\alpha'^2}\left(\frac{1}{\alpha^2} - 1\right) \quad .$$

$$\therefore \quad \frac{\alpha}{\alpha'} = \frac{\alpha}{\sqrt{1 - k^2(1-\alpha^2)}} \quad ,$$

$$\frac{\alpha}{\alpha'} - 1 = \frac{1}{\sqrt{1 + (1-k^2)H}} - 1 \quad . \tag{20.54}$$

Hence, finally
$$\underline{\Gamma_n} = \binom{-\frac{1}{2}}{n} Ny_o\underline{v}(1 - k^2)^n H^n, \; (n = 1, 2, \ldots) \; . \tag{20.55}$$

(d) In section (a) it was mentioned that tertiary and higher terms had not been examined in detail as regards the selection of special grouping of terms. However even if such grouping is not carried out the resulting expressions are still manageable in practice. In particular we may consider tertiary terms in relation to the calculation of tertiary spherical aberration; **i.e.** we shall take $\underline{h}$ to be zero.

Proceeding as above we have, since now $\underline{K}^2 = 0$,

$$\left.\begin{aligned} (1-k)S^{(3)} &= \tfrac{1}{16}(\underline{F}^3 - k\underline{G}^3) \quad , \\ T^{(3)} &= \left[S^{(2)}(F+G) + \tfrac{1}{2}(S^{(1)})^2F\right] + S^{(1)}\left[\tfrac{3}{2}F^2 + 3FG - GH\right] \\ &\quad + \left[\tfrac{5}{16}F^3 + F^2G - \tfrac{1}{2}FH^2 + GH^2 - 2FGH\right] \quad . \end{aligned}\right\} \tag{20.6}$$

Hence by (19.31) we find

$$S^{(3)} = \tfrac{1}{16}\left[5E^3 + 5E^2F - 9E^2H + EF^2 - 3EFH + 3EH^2 + H^3\right] . \tag{20.61}$$

By (19.21) $$J^{(3)} = e_o(S^{(3)} + S^{(2)}T^{(1)} + S^{(1)}T^{(2)}) - u_oT^{(3)} . \qquad (20.62)$$

Substituting in this from (20.11), (20.21), (20.6), (20.61) we obtain after some rearrangement of terms

$$\begin{aligned} J^{(3)} = \; & \tfrac{1}{16}e_o\{5F^2E + 3F^2H + 12\,FE^2 + 10\,FEG + 3\,FEH + 8\,FGH + 3\,FH^2 \\ & + 5\,E^3 + 10\,E^2G - 9\,E^2H - 4\,EGH + 3\,EH^2 - 6\,GH^2 + H^3\} \\ & -\tfrac{1}{8}u_o\{\tfrac{5}{2}F^3 + 7\,F^2E + 8\,F^2G + 6\,F^2H + 4\,FE^2 + 13\,FEG - 4\,FGH \\ & - 4\,FH^2 + 3\,E^2G - 6\,EGH + 3\,GH^2\} . \end{aligned} \qquad (20.7)$$

It should be noted that all but two of the eleven terms in the coefficient of $\underline{u}_o$ already occur in that of $\underline{e}_o$ except for different numerical factors; this implies considerably less work in computing the tertiary spherical aberration than would otherwise be the case.

(e) All equations above contain the tangential aberrations as a special case; we need only put $Z_1 = W_1 = 0$. If this is done we find after some reduction (distinguishing the Γ's in (T 8.5) by bars) that

$$\left.\begin{aligned} \Gamma_1 - \overline{\Gamma}_1 &= 0 \quad , \\ \Gamma_2 - \overline{\Gamma}_2 &= -\tfrac{3}{2}\,\Gamma_1 u^2 \quad ; \end{aligned}\right\} \qquad (20.8)$$

which is as it should be. For in (T § 8) semi-canonical variables were used; and the change to the present canonical variables generates a term

$$\overline{\Gamma}_1\{(1 - \tfrac{1}{2}\tan^2 U)^3 - 1\} = -\tfrac{3}{2}\,\Gamma_1 u^2 + \underline{O}(7) . \qquad (20.81)$$

§21. Iteration in Explicit Form.

(a) For the purpose of practical application it is convenient to write down the explicit form of the higher order coefficients generated by eq. (5.5). The first order coefficients $\underline{a}_p, \ldots, \underline{\bar{c}}_q$ are of course calculated directly by means of (20.41); and then

$$A_{pj} = \sum_{\nu=1}^{j-1} a_{p\nu} = \sum_{\nu=1}^{j-1} \underline{a}_{p\nu} \text{ ; etc.} \tag{21.1}$$

Selecting in (5.5) the terms of the fifth degree we put for the second order coefficients in the expansion of $\Delta \underline{\Lambda}_j$ (omitting the suffix $\underline{j}$)

$$\left.\begin{aligned}
s_1 &= \underline{s}_1 - 3A_q\underline{a} + A_p\underline{\bar{a}} + A_p\underline{b} \\
\bar{s}_1 &= \underline{\bar{s}}_1 - \bar{A}_q\underline{a} + (\bar{A}_p - 2A_q)\underline{\bar{a}} + A_p\underline{\bar{b}} \\
s_2 &= \underline{s}_2 - (2\bar{A}_q + 3B_q)\underline{a} + B_p\underline{\bar{a}} + (\bar{A}_p + B_p - 2A_q)\underline{b} + A_p\underline{\bar{b}} + 2A_p\underline{c} \\
\bar{s}_2 &= \underline{\bar{s}}_2 - \bar{B}_q\underline{a} + (\bar{B}_p - 2\bar{A}_q - 2B_q)\underline{\bar{a}} - \bar{A}_q\underline{b} + (2\bar{A}_p + B_p - A_q)\underline{\bar{b}} + 2A_p\underline{\bar{c}} \\
s_3 &= \underline{s}_3 - 3\bar{C}_q\underline{a} + C_p\underline{\bar{a}} + (C_p - \bar{A}_q)\underline{b} + (2\bar{A}_p - A_q)\underline{c} + A_p\underline{\bar{c}} \\
\bar{s}_3 &= \underline{\bar{s}}_3 - C_q\underline{a} + (\bar{C}_p - 2C_q)\underline{\bar{a}} + (C_p - \bar{A}_q)\underline{\bar{b}} - \bar{A}_q\underline{c} + 3\bar{A}_p\underline{\bar{c}} \\
s_4 &= \underline{s}_4 - 2\bar{B}_q\underline{a} + (\bar{B}_p - 2B_q)\underline{b} + B_p\underline{\bar{b}} + 2B_p\underline{c} \\
\bar{s}_4 &= \underline{\bar{s}}_4 - 2\bar{B}_q\underline{\bar{a}} - \bar{B}_q\underline{b} + (2\bar{B}_p - B_q)\underline{\bar{b}} + 2B_p\underline{\bar{c}} \\
s_5 &= \underline{s}_5 - 2\bar{C}_q\underline{a} + (\bar{C}_p - \bar{B}_q - 2C_q)\underline{b} + C_p\underline{\bar{b}} + (2\bar{B}_p + 2C_p - B_q)\underline{c} + B_p\underline{\bar{c}} \\
\bar{s}_5 &= \underline{\bar{s}}_5 - 2\bar{C}_q\underline{\bar{a}} - \bar{C}_q\underline{b} + (2\bar{C}_p - \bar{B}_q - C_q)\underline{\bar{b}} - \bar{B}_q\underline{c} + (3\bar{B}_p + 2C_p)\underline{\bar{c}} \\
s_6 &= \underline{s}_6 - \bar{C}_q\underline{b} + (2\bar{C}_p - C_q)\underline{c} + C_p\underline{\bar{c}} \\
\bar{s}_6 &= \underline{\bar{s}}_6 - \bar{C}_q\underline{\bar{b}} - \bar{C}_q\underline{c} + 3\bar{C}_p\underline{\bar{c}}
\end{aligned}\right\} \tag{21.2}$$

(If $N_1 \neq 1$ all terms containing primary coefficients are to be divided by N_1). The $\underline{s}_\nu$, $\overline{\underline{s}}_\nu$ are, of course, calculated directly from (20.42). Each of the equations of (21.2) splits up into two equations; *i.e.* those obtained by affixing the subscript *p* or *q* to every term in it; *e.g.*

$$\left.\begin{aligned} s_{1p} &= \underline{s}_{1p} - 3A_q\underline{a}_p + A_p\overline{\underline{a}}_p + A_p\underline{b}_p \\ s_{1q} &= \underline{s}_{1q} - 3A_q\underline{a}_q + A_p\overline{\underline{a}}_q + A_p\underline{b}_q \end{aligned}\right\} \quad (21.21)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

The equations above show exactly how the second order aberration coefficients are compounded of the terms arising at the different surfaces of the system.

In the practical application of (21.2) time may be saved by making use of the first order identities (17.a). For example, considering $\overline{s}_2$ and s_5

$$(\overline{B}_p - 2\overline{A}_q - 2B_q) = -(2\overline{B}_p + 2C_p - B_q) + 4N_1u_pu_q \quad (21.3)$$

(b) There are 84 separate terms in (21.2). Proceeding in the same way with terms of the seventh degree we find that the expressions for t_1, ..., $\overline{t}_{10}$ contain 692 separate terms. (Not all of these need however be computed singly, since certain groups of them recur collectively in different coefficients). Leaving the complete results till a later date we content ourselves here with the expression for t_1, viz.

$$t_1 = \underline{t}_1 - 5A_q\underline{s}_1 + A_p(\overline{\underline{s}}_1 + \underline{s}_2) + 3(A_q^2 - S_{1q})\underline{a} + (S_{1p} - 2A_pA_q)(\overline{\underline{a}} + \underline{b}) + A_p^2(\overline{\underline{b}} + \underline{c}) \quad (21.4)$$

(As before, each coefficient which is represented by a capital letter is understood to be divided by N_1). In (21.4) all the first and second order terms are known from the previous section, remembering that

$$S_{1qj} = \sum_{\nu=1}^{j-1} s_{1q\nu}, \text{ etc.} \tag{21.41}$$

whilst $\underline{t_1}$ is calculated from (20.7).

(c) The equations above do not apply in the case of general linear co-ordinates. To see this we observe that in the case of canonical co-ordinates the coefficient of $\underline{Y_1}\xi$ in $\underline{\delta}_{yj}$ for instance, is $-\frac{1}{N_1} A_q$; to it corresponds the coefficient of $\underline{S}\tilde{\xi}$ in $\underline{\delta}_{sj}$; and by (12.41), (12.72) this is

$$\frac{1}{N_1}\left\{\begin{vmatrix} g_p & g_q \\ \pi & \rho \end{vmatrix} F^{(1)'}_{00,k|a} + \left(\begin{vmatrix} \bar{g}_p & \bar{g}_q \\ \pi & \rho \end{vmatrix} + g\right) F^{(1)'}_{00,k|b} - g\, F^{(1)}_{00,j|b}\right\} . \tag{21.5}$$

However, (21.5) indicates how in this case the required equations corresponding to (21.2) can be written down by inspection of the latter. And since the first two terms in (21.5) do not vary from surface to surface the increase in complexity is far smaller than might be expected at first sight.

§22. Jacoby Polynomials.

Various expansions may be obtained in compact form by means of the so-called Jacoby-Polynomials. (Madelung, 1943). Since they seem to occur frequently in the algebraic theories of optics we consider them briefly. The polynomial of order $\underline{n}$ is defined by

$$J_n(p; q; x) = F(-n, p+n; q; x) , \tag{22.1}$$

where F denotes the hypergeometric function of Gauss. (loc.cit.p.184)

Hence

$$J_n\ (p;\ q;\ x) \quad = \quad \sum_{\nu=0}^{\infty}(-1)^{\nu}\binom{n}{\nu}\frac{\Gamma(p+n+\nu)\Gamma(q)}{\Gamma(p+n)\Gamma(q+\nu)}\ x^{\nu} \quad . \tag{22.11}$$

Thus we get for the polynomials of (T § 6)

$$g_n(x) \quad = \quad (-1)^n\binom{\frac{1}{2}}{n}\,J_n(-n-\tfrac{1}{2};\ \ -n+\tfrac{3}{2};\ x)\ . \tag{22.2}$$

It is convenient to define

$$\widetilde{M}_n(x) \quad = \quad J_n(\tfrac{1}{2}-n;\ \tfrac{3}{2}-n;\ x) \quad , \tag{22.3}$$

$$M_n(x) \quad = \quad J_{[n]}(\tfrac{1}{2}-n;\ \tfrac{3}{2}-n;\ x) \quad , \tag{22.4}$$

$$\text{where} \qquad [n] \quad = \quad \begin{cases} \tfrac{1}{2}n, & (n \text{ even}) \\ \tfrac{1}{2}(n-1), & (n \text{ odd}) \end{cases} \quad ; \tag{22.41}$$

for then we have the expansions

$$(1 + t + \tfrac{1}{4}\,x\,t^2)^{\frac{1}{2}} \quad = \quad \sum_{n=0}^{\infty} M_n(x)t^n \tag{22.5}$$

$$(1 + t)^{\frac{1}{2}}(1 + \ x\ t)^{-\frac{1}{2}} \quad = \sum_{n=0}^{\infty} \widetilde{M}_n(x)t^n \quad . \tag{22.6}$$

For the first few of the polynomials we find

$$\left.\begin{array}{llll} M_0(x) = 1 & & \widetilde{M}_0(x) = 1 \\ M_1(x) = \frac{1}{2} & & \widetilde{M}_1(x) = \frac{1}{2}(1-x) \\ M_2(x) = -\frac{1}{8}(1-x) & & \widetilde{M}_2(x) = -\frac{1}{8}(1-x)(1+3x) \\ M_3(x) = +\frac{1}{16}(1-x) & & \widetilde{M}_3(x) = +\frac{1}{16}(1-x)(1+2x+5x^2) \\ M_4(x) = -\frac{1}{128}(1-x)(5-x) & & \widetilde{M}_4(x) = -\frac{1}{128}(1-x)(5+9x+15x^2+35x^3) \\ M_5(x) = +\frac{1}{256}(1-x)(7-3x) & & \widetilde{M}_5(x) = +\frac{1}{256}(1-x)(7+12x+18x^2+28x^3+63x^4) \end{array}\right\} \tag{22.7}$$

(22.5) is a special case of the more general development

$$(1 + t + \tfrac{1}{4} x t^2)^s = \sum_{n=0}^{\infty} \binom{s}{n} J_{[n]}(\tfrac{1}{2} - m; \quad s - n + 1; \quad x) t^n \quad . \tag{22.8}$$

If $\underline{s}$ is a positive integer we use

$$\binom{s}{n} J_{[n]}(\tfrac{1}{2} - m; \quad s - m + 1; \quad x) = \sum_{\nu=0}^{[n]} \binom{n-\nu}{\nu}\binom{s}{n-\nu}\left(\frac{x}{4}\right)^{\nu} \quad . \tag{22.81}$$

Eq. (22.8) was already referred to in §8.b.

§23. First Method. Lower Orders in Semi-Canonical Variables.

(a) The expansion of $\underline{\Delta}$ may be obtained *ab initio* in the same way as the expansion for $\underline{J}$ was obtained in §20. Alternatively, once the latter is known the former may easily be derived from it. Thus by means of (14.6) we may rewrite (15.6) as

$$\underline{\Delta} = \underline{\beta}\left(\frac{\alpha'}{\alpha} - 1\right) - \varpi \underline{C}^* \left(\frac{1}{\sigma\alpha} - 1\right) \quad , \tag{23.1}$$

$$\text{or } \underline{\Delta} = - \frac{\underline{\beta}\,\underline{T}}{1 + \underline{T}} - \varpi \underline{C}^* \underline{S} \quad . \tag{23.11}$$

To develop (23.11) we need only write

$$E = \frac{E^*}{\alpha^2} = \frac{E^*}{1 - H^*}; \quad \text{etc.} \tag{23.2}$$

so that

$$E^* = k(\sin^2 I_y^* + \sin^2 I_z^*)$$

$$\rightarrow i^* i^{*\prime} \equiv i_p i_p{}' \xi^* + 2 i_p i_q{}' \eta^* + i_q i_q{}' \zeta^* \quad , \tag{23.21}$$

and F^*, G^*, H^* similarly. It is scarcely necessary again to go through the resulting substitutions; we shall merely quote the final result. If we write

$$\left.\begin{aligned}
\gamma_1^* &= \tfrac{1}{2}(i^* i^{*\prime} + u^{*2}) \\
\gamma_2^* &= \tfrac{1}{2}(u^{*2} - u^{*\prime 2}) \\
\gamma_3^* &= -e^* u^{*\prime} \\
\gamma_4^* &= \tfrac{1}{4}(u^{*\prime 2} - 4u^* u^{*\prime} + 3u^{*2}) \quad ;
\end{aligned}\right\} \tag{23.3}$$

then
$$\left.\begin{aligned}
\underline{\Delta}^{(1)} &= \underline{e}^* \gamma_1^* + \underline{v}^* \gamma_2^* \\
\underline{\Delta}^{(2)} &= \underline{e}^* (\tfrac{3}{2}\gamma_1^{*2} - \cancel{k}\gamma_2^* \gamma_4^*) + \underline{v}^* (\gamma_1^* \gamma_3^* + \gamma_2^* \gamma_4^*) \quad .
\end{aligned}\right\} \tag{23.4}$$

(b) The <u>plane surface</u> is again easily dealt with. Since now

$$\Delta Nu_o = \Delta N\underline{\beta} = \Delta N\underline{Y} = 0 \quad , \tag{23.5}$$

we have
$$\Delta\underline{\Lambda}^* = - Nu_o \underline{Y}^* (\frac{\alpha'}{\alpha} - 1) \quad . \tag{23.51}$$

But by (20.54)
$$\frac{\alpha'}{\alpha} - 1 = \sqrt{\frac{1 - k^2 H^*}{1 - H^*}} \quad . \tag{23.52}$$

$$\left.\begin{aligned}
\therefore \ (\Delta\underline{\Lambda}^*)^{(1)} &= -\tfrac{1}{2} Nu_o (1 - k^2) \underline{y}^* H^* \\
(\Delta\underline{\Lambda}^*)^{(2)} &= -\tfrac{1}{8} Nu_o (1 - k^2)(3 + k^2) \underline{y}^* H^{*2} \quad .
\end{aligned}\right\} \tag{23.6}$$

By (22.6) we may write quite generally

$$(\Delta\underline{\Lambda}^*)^{(n)} = (-1)^{n+1} Nu_o k^{2n} \widetilde{M}_n (1/k^2) \underline{y}^* H^{*n} \quad . \tag{23.7}$$

(c) For the purpose of iteration we may use all the relevant equations developed previously in § 21. The only change that occurs is purely formal, <u>i.e.</u> all symbols used now appear with an asterisk.

§24. Second Method. All Orders.

(a) The principle of this alternative method was explained in §18.a. In this paragraph it will be considered explicitly for the case of canonical co-ordinates. Since we now allow the presence of accented variables probably the most convenient expression for $\Delta\underline{\Lambda}$ is

$$\Delta\underline{\Lambda} = e_o\underline{C}'\cdot\left(\frac{1}{\alpha\sigma} - 1\right) - u_o\underline{C}\left(\frac{\alpha}{\alpha'} - 1\right)$$

$$= e_o\underline{C}'\underline{S} - u_o\underline{CT} \quad . \tag{24.1}$$

For $\underline{S}$ we may use the expression (19.4). Now the terms of degree $\underline{2n}$ in the expansion of the first square root appear as the coefficient of $\underline{t}^n$ in the expansion of

$$\sqrt{1 + \underline{F}t + k^2\underline{K}^2t^2} \quad .$$

By (22.5) this is $\underline{F}^n M_n(4k^2\underline{K}^2/\underline{F}^2)$. Treating the other root in like manner we get

$$(1-k)S^{(n)} = \underline{F}^n M_n(4k^2\underline{K}^2/\underline{F}^2) - k\underline{G}^n M_n(4\underline{K}^2/\underline{G}^2) \quad . \tag{24.2}$$

Also $\quad \underline{T} \mp \sqrt{\dfrac{1 + H'}{1 + H}} - 1 \;;$

therefore, by (22.6)

$$T^{(n)} = \bar{H}'^n \tilde{M}_n(H/\bar{H}') \, . \tag{24.3}$$

Hence finally, for $n = 1, 2, 3, \ldots$

$$(\Delta\underline{\Lambda})^{(n)} = i_o\underline{\bar{c}}\{\underline{F}^n M_n(4k^2\underline{K}^2/\underline{F}^2) - k\underline{G}^n M_n(4\underline{K}^2/\underline{G}^2)\} - u_o\underline{c}\bar{H}'^n\tilde{M}_n(H/\bar{H}') \, . \tag{24.4}$$

Notice that for tangential rays the only surviving terms in $S^{(n)}$ are

$$\binom{\frac{1}{2}}{n}(\underline{F}^n - k\underline{G}^n) \quad . \tag{24.41}$$

The plane surface may again be dealt with by means of §20.c.

(b) For the purpose of iteration expressions analogous to (21.2) could be written down.

But owing to the distinction between barred and unbarred variables we are now effectively dealing with homogeneous polynomials in eight variables instead of only four. In order therefore to avoid the appearance of a multiplicity of symbols or affixes it is best to leave the writing down of the terms due to the increments until the current symbols (v.§25.b) in any particular computation have become available. It is then a reasonably simple matter, not very different from the considerations of §21.

§25. Sample Computation.

(a) For the purpose of demonstrating the practical application of the methods developed above, all the computations necessary for the determination of the complete primary and secondary aberrations and of the tertiary spherical aberration of a corrected Cooke Triplet have been carried out. The system used is the same as that of (T§15.) and the object is again assumed to be at infinity. To eliminate the appearance of a vast number of noughts in the calculations the dimensions of the system have been reduced so as to make the focal length of the system exactly equal to unity. Its specifications are therefore

Surface	1	2	3	4	5	6
r	+0.2072802	-1.326386	-0.6078840	+0.1953913	+3.218270	-0.6843864
N'	1.6162	1.0	1.5725	1.0	1.6162	1.0
d'	.0402775	.0168512	.0096145	.1387382	.0313246	-

($N_1 = 1$; $f' = 1$; $d = .041353$, behind 4th surface.

Max. stop-number of system $= f'/5.6$) (25.1)

Since $u_{\alpha 1} = 0$ and $f' = 1$ the factor μ' is unity and can therefore be ignored.

(b) TABLE I.

This contains in detail all the calculations necessary for the determination of the full primary and secondary aberrations of the system. The final aberration coefficients appear in the seventh column, marked Σ_1^k. The entries in the corresponding rows constitute the <u>contributions</u> by the surface to that coefficient. For example, considering S_{1k}' (which, since $u_{\alpha 1} = 0$, here $= S_{1pk}'$ and represents the spherical aberration) the row marked t_{123} gives the individual contributions by the surfaces to the spherical aberration, the latter itself appearing in the seventh column. Notice also that if the part of the system preceding the jth surface behaved as a Gaussian System then t_{127} would be the only surviving entry at the jth surface contributing to t_{123}. In this way the secondary contributions to any coefficient by a given surface may themselves be sub-

TABLE I.

	$j \longrightarrow$		1	2	3	4	5	6
p-trace	$y_- - d'u = y$		$\underline{1}$	·925916	·867871	·862766	·840793	·836000
	$u'_- = u$	t_1	$\underline{0}$	1·33937	3·40294	1·64383	·0863062	·153008
	$y/r - u = i$		4·82439	−2·55744	−4·93178	2·72057	·174950	−1·37454
	$k\,i = i'$		2·98502	−4·10101	−3·07268	4·27810	·108248	−2·22158
	$y/r - i' = u'$	t_2	1·33937	3·40294	1·64383	·0863062	·153008	1·00000
q-trace	$y_- - d'u = y$		$\underline{0}$	−·024921	−·0₃14772	−·0478600	−·209880	−·231735
	$u'_- = u$	t_3	$\underline{1}$	·618735	·988422	·653469	1·16781	·697701
	$y/r - u = i$		−1·00000	−·599947	−·920025	−·898413	−1·23303	−·859099
	$k\,i = i'$		−·618735	−·969634	−·585072	−1·41275	−·762916	−·580375
	$y/r - i' = u'$	t_4	·618735	·988422	·653469	1·16781	·697701	·918978
	c_p	t_5	1·00000	5·43953	2·93716	·885903	·563036	1·52039
	$i_p - i'_p = e_p$	t_6	1·83937	1·56367	−1·75911	−1·55753	·0667022	·846992
	c_q	t_7	−·207280	1·28611	5·59269	−·276040	−3·96321	·397201
	$i_q - i'_q = e_q$	t_8	−·381265	·369687	−·334954	·514841	−·470109	·221277
	$i_p i'_p$	t_9	14·4009	10·4061	14·8465	11·6389	·0189379	3·05358
	$i_p i'_q$	t_{10}	−2·98502	2·46039	2·82694	−3·84350	−·133472	·797748
	$i_q i'_q$	t_{11}	·618735	·581723	·538281	1·26924	·940695	·208412
	u_p^2	t_{12}	0	3·38328	11·5800	2·70219	·00744876	·0234116
	$u_p u'_p$	t_{13}	0	6·25026	5·59387	·141873	·0132056	·153008
	u'^2_p	t_{14}	3·38328	11·5800	2·70219	·00744876	·0234116	1·00000
	u_q^2	t_{15}	1·00000	·382833	·976979	·427021	1·36378	·486787
	$u_q u'_q$	t_{16}	·618735	·611672	·645903	·763127	·814782	·641172
	u'^2_q	t_{17}	·382833	·976979	·427021	1·36378	·486787	·844520
	$u_p u'_q$		0	1·91707	2·22371	1·91969	·0602159	·140611
	$u_q u'_p$		1·83937	2·10552	1·62480	·0563984	·178685	·697701
	$u_p u'_q + u_q u'_p$	t_{18}	1·83937	3·92359	3·84852	1·97608	·238903	·838812
	$u_p u_q$	t_{19}	0	1·13808	3·36354	1·07419	·100789	·106754
	$u'_p u'_q$	t_{20}	1·13808	3·36354	1·07419	·100789	·106754	·918978
	$K = k/(1-k)^2$	t_{21}	4·25649	4·25649	4·79777	4·79777	4·25649	4·25649
γ_1	$t_9 + t_{14}$	t_{22}	17·7842	21·9861	17·5487	11·6463	·0423495	4·05358
	$t_{10} + t_{20}$	t_{23}	−1·84694	5·82393	3·90112	−3·74271	−·0267180	1·71673
	$t_{11} + t_{17}$	t_{24}	1·00157	1·55871	·965303	2·63302	1·42748	1·05293
γ_2	$t_{12} - t_{14}$	t_{25}	−3·38328	−8·19672	8·87781	2·69474	−·0159628	−·976088
	$t_{19} - t_{20}$	t_{26}	−1·13808	−2·22546	2·28935	·973404	−·0059648	−·812224
	$t_{15} - t_{17}$	t_{27}	·617167	−·594144	·549957	−·936759	·876993	−·357733
$e_p\gamma_1$	$t_6 t_{22}$	t_{28}	32·7117	34·3768	−30·8700	−18·1395	·00282490	3·43335
	$t_6 t_{23}$	t_{29}	−3·39720	9·10613	−6·86251	5·82937	−·00178215	1·45405
	$t_6 t_{24}$	t_{30}	1·84225	2·43715	−1·69807	−4·10100	·0952162	·891824
$e_q\gamma_1$	$t_8 t_{22}$	t_{31}	−6·78048	8·12797	−5·87800	5·99019	−·0199089	·891963
	$t_8 t_{23}$	t_{32}	·704172	2·15303	−1·30670	−1·92503	·0125604	·379871
	$t_8 t_{24}$	t_{33}	−·381863	·576234	−·323332	1·35427	−·671072	·232989
$u'_p\gamma_2$	$t_2 t_{25}$	t_{34}	−6·22310	−27·8929	14·5936	·232573	−·00244244	−·976588
	$t_2 t_{26}$	t_{35}	−2·09336	−7·57310	3·76331	·0840108	−·0₃91266	−·812224
	$t_2 t_{27}$	t_{36}	1·13520	−2·02184	·904037	−·080848	·134187	−·357733
$u'_q\gamma_2$	$t_4 t_{25}$	t_{37}	−2·09335	−8·10182	5·80137	3·14694	−·0111373	−·897463
	$t_4 t_{26}$	t_{38}	−·704172	−2·19969	1·49602	1·13675	−·0041642	−·746415
	$t_4 t_{27}$	t_{39}	·381863	−·587266	·359379	−1·09396	·611879	−·328749
$e_p\gamma_1 + u'_p\gamma_2$	$\frac{1}{2}(t_{28} + t_{34})$	t_{40}	13·2443	3·24194	−8·13819	−8·95345	·0₃191180	1·22838
	$t_{29} + t_{35}$	t_{41}	−5·49056	1·53303	−3·09921	5·91338	−·00269481	·641829
	$\frac{1}{2}(t_{30} + t_{36})$	t_{42}	1·48873	·207656	−·397017	−2·09092	·114702	·267046
$e_q\gamma_1 + u'_q\gamma_2$	$\frac{1}{2}(t_{31} + t_{37})$	t_{43}	−4·43692	·013073	−·038314	4·56866	−·0155231	−·0₃2500
	$t_{32} + t_{38}$	t_{44}	0	−·046662	·189319	−·788279	·0083982	−·366544
	$\frac{1}{2}(t_{33} + t_{39})$	t_{45}	0	−·055160	·0180238	·130157	−·0295964	−·0478798

(Table I contd.)

	$j \longrightarrow$		1	2	3	4	5	6	Σ_1^k
	$a_p = t_5 t_{40}$	t_{46}	13.2443	17.6846	-23.9032	-7.49421	$.0_{3}107641$	1.86762	
	$b_p = t_5 t_{41}$	t_{47}	-5.49055	8.33977	-9.40287	4.94301	-.00151727	.975929	
	$c_p = t_5 t_{42}$	t_{48}	1.49873	1.12954	-1.16610	-1.74781	.0646813	.406012	
	$a_q = t_5 t_{43}$	t_{49}	-4.43692	.071111	-.112530	3.91888	-.00874006	$-.0_{3}3801$	
	$b_q = t_5 t_{44}$	t_{50}	0	-2.55514	.556061	-.658925	.0047295	-.557289	
	$c_q = t_5 t_{45}$	t_{51}	0	-.0300058	.0629387	.108799	-.0166638	-.0727968	
	$\bar{a}_p = t_7 t_{40}$	t_{52}	-2.74528	4.16948	-4.55144	2.47151	$-.0_{3}78642$	.487914	
	$\bar{b}_p = t_7 t_{41}$	t_{53}	1.13808	1.97165	-1.73329	-1.63233	.0106986	.254935	
	$\bar{c}_p = t_7 t_{42}$	t_{54}	-.308583	.267066	-.222089	.577178	-.456160	.106071	
	$\bar{a}_q = t_7 t_{43}$	t_{55}	.919695	.016613	-.021429	-1.26111	.065955	$-.0_{4}9980$	
	$\bar{b}_q = t_7 t_{44}$	t_{56}	0	-.060012	.105990	.217596	-.033326	-.145592	
	$\bar{c}_q = t_7 t_{45}$	t_{57}	0	-.0070942	.010[illegible]	-.0369285	.117445	-.0190179	
	$A_p = \Sigma_1^{j-1} t_{46}$	t_{58}	0	13.2443	30.8759	6.97569	-.509522	-.508414	1.86920
	$B_p = \Sigma_1^{j-1} t_{47}$	t_{59}	0	-5.49055	2.94942	-6.25446	-1.31144	-1.31290	-.337183
	$C_p = \Sigma_1^{j-1} t_{48}$	t_{60}	0	1.49873	2.61827	1.45217	-.295642	-.221061	.174952
	$A_q = \Sigma_1^{j-1} t_{49}$	t_{61}	0	-4.43692	-4.36580	-4.47933	-.659457	-.668197	
	$B_q = \Sigma_1^{j-1} t_{50}$	t_{62}	0	0	-.255514	.302247	-.356678	-.351949	
	$C_q = \Sigma_1^{j-1} t_{51}$	t_{63}	0	0	-.0300058	.0229331	.131732	.115065	
	$\bar{A}_p = \Sigma_1^{j-1} t_{52}$	t_{64}	0	-2.74528	1.42421	-3.12723	-.655722	-.656481	-.165567
	$\bar{B}_p = \Sigma_1^{j-1} t_{53}$	t_{65}	0	1.13808	3.10973	1.37644	-.256790	-.245196	.009739
	$\bar{C}_p = \Sigma_1^{j-1} t_{54}$	t_{66}	0	-.308583	-.045170	-.263566	.313622	-.14539	-.0354652
	$\bar{A}_q = \Sigma_1^{j-1} t_{55}$	t_{67}	0	.919696	.936498	.915070	-.346036	-.284437	
	$\bar{B}_q = \Sigma_1^{j-1} t_{56}$	t_{68}	0	0	-.060012	.046868	.263464	.230138	
	$\bar{C}_q = \Sigma_1^{j-1} t_{57}$	t_{69}	0	0	-.0070942	.0029860	-.032943	.084602	
	$-\frac{1}{4}(3t_{12}+t_{14})$	t_{70}	-.845920	-5.43246	-9.36055	-2.02950	-.0114895	-.267559	
	$-\frac{1}{2}(3t_{19}+t_{20})$	t_{71}	-.569042	-3.38890	-5.59241	-1.66161	-.204561	-.619620	
	$-\frac{1}{4}(3t_{18}+t_{17})$	t_{72}	-.845708	-.331370	-.839489	-.661211	-1.14453	-.576220	
γ_3	$t_{70}+\frac{3}{4}t_{22}$	t_{73}	12.4923	11.0571	5.80097	6.70624	.0203226	2.77263	
	$t_{71}+\frac{1}{2}t_{23}$	t_{74}	-3.33945	5.34700	.269271	-7.27575	-.244638	1.95547	
	$t_{72}+\frac{3}{4}t_{24}$	t_{75}	-.0945318	.637660	-.115513	1.31355	-.073921	.213479	
γ_4	$t_{70}+t_{15}$	t_{76}	-.845920	.926803	-3.76668	-1.88663	.00176611	-.114660	
	$t_{71}+t_{18}$	t_{77}	1.27033	.534697	-1.73390	.314399	.0243423	.219692	
	$t_{72}+t_{16}$	t_{78}	-.226973	.0802020	-.193686	.101916	-.329750	.0649616	
	$t_{21} t_6$	t_{79}	7.82926	6.65533	-8.43980	-7.47266	.253917	8.60521	
$(\kappa_p\gamma_1+\kappa_q\gamma_2)$	$t_1 t_{22}$		0	40.4405	59.7171	19.1446	.00365502	.620232	
	$t_{79} t_{25}$		-26.48[illegible]	-54.5519	-74.9269	-20.1369	-.00453212	-3.52081	
	$\frac{1}{2}\Sigma$	t_{80}	-13.2442	-7.05569	-7.60489	-.49613	$-.0_{4}38548$	-1.45029	
	$t_1 t_{23}$		0	10.7236	13.27532	-6.15239	-.00230693	.262674	
	$t_{79} t_{26}$		-5.91035	-14.8112	-19.3216	-7.27392	-.0016935	-2.92824	
	Σ	t_{81}	-5.91035	-4.09880	-6.04632	-13.42631	-.0039994	-2.66556	
	$t_1 t_{24}$		0	2.86704	3.28487	4.32824	.123201	.161107	
	$t_{79} t_{27}$		4.83196	-3.95423	-4.64152	7.00008	.248994	-1.28970	
	$\frac{1}{2}\Sigma$	t_{82}	2.41598	-.543596	-.678327	5.66416	.186097	-.564298	
	$t_{40} t_{73}$		165.452	36.8465	-30.9330	-60.0439	$.0_{3}388528$	3.40585	
	$t_{76} t_{80}$		11.2023	-5.83367	28.6452	.98601	-.0774524	.166131	
	$\delta_1 = \Sigma$	t_{83}	176.654	30.0128	-2.2879	-59.1079	$.0_{3}311076$	3.57198	
	$t_{41} t_{73}$		-69.6997	16.9509	-11.7800	39.6565	$-.0_{4}54765$	1.77955	
	$t_{40} t_{74}$		-44.2286	17.3346	-2.19154	65.1430	$-.0_{4}467699$	2.40206	
	$t_{76} t_{81}$		7.53654	-3.38890	22.7746	25.3305	$-.0_{3}706338$	.805341	
	$t_{77} t_{80}$		-16.8246	-3.77266	13.1861	-.15698	$-.0_{4}150607$	.317167	
	$\delta_2 = \Sigma$	t_{84}	-122.106	27.1240	21.9892	129.974	$-.0_{3}123660$	4.16979	

(Table I contd.)

	$J \longrightarrow$		1	2	3	4	5	6	Σ_i^k
$J^{(2)}$	$t_{62}t_{70}$		18.5976	2.29807	1.60905	-14.0222	00233105	740418	
	$t_{60}t_{75}$		-1.25201	2.06726	940063	-11.7608	$-0_4 14182$	262233	
	$t_{75}t_{82}$		-3.04348	-449447	2.55501	-10.6862	$-0_3 328668$	0646405	
	$t_{78}t_{80}$		3.00609	-516850	1.47222	-.050564	$0_3 144611$	-0941985	
	$\hat{s}_1 = \Sigma$	t_{84}	18.2082	3.34801	3.65827	-36.5198	00279020	973094	
	$t_{61}t_{74}$		18.2354	8.19712	-834691	-43.0243	$0_3 659253$	1.25508	
	$t_{77}t_{81}$		-11.2191	-2.19162	10.4827	-4.22122	$-0_3 13735$	-582939	
	$\hat{s}_2 = \Sigma$	t_{86}	7.01635	6.00550	9.6491	-47.2455	$0_3 52$[illegible]	67214	
	$t_{63}t_{76}$		-4.97152	1.11034	-106913	15.21302	-0280605	622199	
	$t_{61}t_{78}$		519032	977653	357997	7.767033	$0_3 19920$	137017	
	$t_{77}t_{82}$		3.06908	-290639	1.17615	1.78081	0063100	-123408	
	$t_{78}t_{81}$		2.02241	-328732	1.17049	-1.36836	0012168	-173133	
	$\hat{s}_{60} = \Sigma$	t_{87}	639003	1.46850	2.59773	23.3930	-0201515	362675	
	$t_{62}t_{75}$		-140732	132414	0458604	-2.74654	-0084789	0570085	
	$t_{77}t_{82}$		-548362	-0435978	131316	577270	-0613655	-0366521	
	$\hat{s}_4 = \Sigma$	t_{88}	-689094	088816	177175	-2.16927	-0698444	0203664	
	$t_{62} - t_{64}$	t_{89}	0	6.12855	10.15679	5.82942	663171	679893	
	$2t_{61} + 3t_{62}$	t_{90}	0	1.83937	1.11155	2.73688	-1.76211	-1.62472	
	$t_{89} + t_{59}$	t_{91}	0	628001	13.00421	-.42504	-648273	-633068	
	$t_{19} - t_{62}$	t_{92}	0	1.10765	3.61736	771946	457467	458703	
	$t_{92} - 2t_{67}$	t_{93}	0	-.701286	1.74436	-1.05819	1.14954	1.02758	
	$t_{12} + t_{61}$	t_{94}	0	-1.05364	7.21420	-1.77615	-652008	-644755	
	$t_{94} + t_{59}$	t_{95}	0	-6.54419	10.0626	-8.03061	-1.96345	-1.95775	
	$t_{66} - t_{68}$	t_{96}	0	-.308583	018495	-.309424	050158	-271677	
	$t_{19} + t_{66}$	t_{97}	0	2.27617	6.47327	2.45063	-155101	-138442	
	$t_{96} - t_{68}$	t_{98}	0	-308583	078507	-.355292	-213306	-601816	
	$t_{97} + 2t_{60}$	t_{99}	0	5.25362	11.70981	5.35496	-746385	-600564	
	$2t_{66} - t_{65}$	t_{100}	0	-617167	-053028	-550046	495511	-398146	
	$t_{100} - t_{65}$	t_{101}	0	-617167	006984	-595914	232047	-628284	
	$4t_{89} + t_{90}$	t_{102}	0	6.89170	14.5657	7.03363	-1.36895	-1.19771	
	$-3t_{61}t_{46}$		0	234.730	-215.070	-100.550	$0_3 212104$	8.74380	
	$t_{65}t_{52}$		0	66.2218	-140.565	17.2405	$0_3 386786$	-24806	
	$t\ t_{47}$		0	110.444	-281.087	34.4809	$0_3 77$[illegible]	-49[illegible]	
	$\underline{s}_1 = t_5 t_{33}$	(t_{127})	176.654	163.256	-6.7[illegible]	-41.4085	081751	5.43079	
	$s_1 = \Sigma$	(t_{123})	176.654	563.651	-741.420	-98.2375	0013721	8.43040	-90.9206
	$-t_{67}t_{46}$		0	-14.2153	22.3863	6.54958	$0_4 372477$	581219	
	$t_{89}t_{62}$		0	35.9521	-46.5237	14.4075	$-0_3 502109$	331730	
	$t_{58}t_{53}$	t_{103}	0	26.1130	-53.5220	-11.3866	$-005_4 37793$	-129613	
	$\underline{\bar{s}}_1 = t_7 t_{33}$	(t_{128})	-86.6169	38.5917	-1.2796	16.3161	$-0_4 12344$	1.41879	
	$\bar{s}_1 = \Sigma$		-86.6169	74.0475	-73.6396	26.1956	-005916	2.15213	-12.8773
	$-t_{90}t_{46}$		0	+32.4366	26.5676	20.4834	$0_3 189675$	3.03435	
	$t_{59}t_{62}$		0	-22.5929	-12.9644	-15.4579	$0_3 994916$	-640612	
	$t_{91}t_{47}$		0	6.32036	-119.370	-2.10098	$0_3 983605$	-617765	
	t_{103}		0	26.1130	-58.5220	-11.3866	-00543743	-129613	
	$2t_{58}t_{48}$		0	29.4200	-72.0159	-24.3843	-0656920	-412845	
	$\underline{s}_2 = t_5 t_{34}$	(t_{129})	-122.106	147.542	64.5859	108.646	$-0_4 69625$	6.83969	
	$s_2 = \Sigma$		-122.106	163.366	-165.723	75.7493	-.069021	7.57321	-50.9598
	$-t_{68}t_{46}$	t_{104}	0	0	-1.4345	34329	$-0_4 283596$	-429809	
	$t_{93}t_{52}$		0	-2.92400	-7.9393	-2.61533	$-0_3 872089$	501370	
	$-t_{67}t_{47}$		0	-7.66122	8.5248	-4.52320	$-0_3 525030$	277562	
	$t_{95}t_{53}$		0	-12.9028	-17.444	13.1086	-0209764	-499098	
	$2t_{58}t_{54}$	t_{105}	0	7.07420	-13.7127	8.05243	462918	-107850	
	$\underline{\bar{s}}_2 = t_7 t_{34}$		25.3102	34.3843	12.2979	-35.8780	$0_3 490707$	1.66624	
	$\bar{s}_2 = \Sigma$		25.3102	18.4625	-19.7052	-21.5123	44098	1.39841	4.3946

(Table I contd.)

$j \longrightarrow$		1	2	3	4	5	6	Σ^h
$-3\,t_{45}t_{46}$		0	0	-2.15170	.514909	$-0_4425893$	-.644767	
$t_{46}t_{52}$		0	6.20722	-11.9169	3.59904	$0_3224286$	-.112738	
$\frac{1}{2}t_{49}t_{47}$		0	4.74521	-15.2030	2.66488	$-0_4764623$	.0620868	
$t_{44}t_{48}$		0	-1.17013	-7.41260	3.10436	-.0421075	-.261791	
$\frac{1}{2}t_{105}$		0	5.53710	-6.55633	4.02621	.221459	-.0539275	
$s_3 = t_5 t_{95}$		15.3092	15.2116	10.1675	-30.5270	.0016787	1.47948	
$s_3 = \Sigma$		7.3092	31.5110	-34.4989	-16.6376	.191027	.458401	-.66780
$-t_{47}t_{46}$	t_{106}	0	0	-.16967	.022348	0_335460	-.16782	
$t_{46}t_{52}$		0	-1.25663	-.05413	-.764744	$-0_4380520$	-.181347	
$\frac{1}{2}t_{49}t_{53}$		0	1.12196	-2.71500	-.476715	$0_3538900$	.0136076	
$-t_{67}t_{45}$		0	-1.03842	1.49205	1.59987	.0223475	.115455	
$2\,t_{44}t_{54}$		0	-2.19951	-.945070	-5.41490	.895376	-.208900	
$\bar{s}_3 = t_7 t_{36}$		-3.79494	4.30591	1.93410	10.0909	-.0110721	.386514	
$\bar{s}_3 = \Sigma$		-3.79494	.90289	-1.09129	4.64626	.907157	-.03246	.153762
$2\,t_{104}$		0	0	2.7690	.68658	$-0_4567190$	-.569618	
$t_{92}t_{47}$		0	9.49044	-32.9283	3.81574	$-0_3694101$	.447616	
$t_{54}t_{53}$		0	-10.5254	-4.93713	10.2093	-.0140241	-.334720	
$2\,t_{49}t_{45}$		0	-12.4036	-6.64307	21.8632	-.169390	-1.06616	
$s_4 = t_5 t_{86}$		7.01635	32.6671	28.3410	-39.4927	0_329386	1.0219	
$s_4 = \Sigma$		7.01635	18.9285	-19.0365	-2.91782	-.153871	-.7910	3.0157
$-2\,t_{65}t_{52}$		0	0	-.540282	-.22673	$0_3399750$	-.224575	
$-t_{65}t_{47}$		0	0	-.546282	-.22673	$0_3399750$	-.224575	
$t_{97}t_{53}$		0	4.48779	-11.2201	-4.00024	-.00165867	-.0362137	
$2\,t_{59}t_{54}$	t_{107}	0	-2.93268	-1.26492	-7.21987	1.19384	-.278533	
$\bar{s}_4 = t_7 t_{90}$		-1.45435	7.72373	5.3964	13.0416	-.0020710	.20477	
$\bar{s}_4 = \Sigma$		-1.45435	9.27884	-8.18113	1.36807	1.19091	-.49593	1.70641
$2\,t_{106}$		0	0	-.33914	.044676	0_570920	-.31564	
$t_{98}t_{47}$		0	-2.57327	-.71464	-1.75621	$0_3323643$	-.587268	
$t_{60}t_{53}$		0	2.93524	-4.53822	-2.37041	-.00316148	-.0589056	
$t_{99}t_{48}$		0	5.93419	-13.6548	-7.35945	-.0482025	-.243836	
$\frac{1}{2}t_{107}$		0	-1.46634	-.632460	-3.60994	.696917	-.139267	
$s_5 = t_5 t_{87}$		.639003	7.98794	7.62996	19.5643	-.0113460	.551406	
$s_5 = \Sigma$		.639003	12.8178	-12.2493	2.50297	.534538	-.79351	3.45142
$-2\,t_{69}t_{52}$		0	0	-.064578	-.014760	-0_449984	-.082459	
$-t_{69}t_{47}$	t_{108}	0	0	-.064578	-.014760	-0_449984	-.082459	
$t_{101}t_{53}$		0	-1.21683	-.01211	.972728	.00248142	-.160172	
$-t_{68}t_{48}$		0	0	-.069980	.080167	-.0170148	-.0934389	
$t_{102}t_{54}$		0	1.70701	-3.23416	4.05967	.619540	-.127041	
$\bar{s}_5 = t_7 t_{87}$		-.132453	1.88865	1.45283	-6.45740	.0799653	.144065	
$\bar{s}_5 = \Sigma$		-.132453	2.37882	-1.99258	-1.37436	.683871	-.401516	-.838213
t_{108}		0	0	-.064578	-.014760	-0_449984	-.082459	
$t_{100}t_{48}$		0	-.697116	.061836	.961375	.0320008	-.161652	
$t_{60}t_{54}$		0	.397589	-.581359	.838129	.134565	-.0245088	
$s_6 = t_5 t_{88}$		-.689094	.48312	.520392	-1.81330	-.0393249	.0309496	
$s_6 = \Sigma$		-.689094	.18359	-.063709	-.02866	.127191	-.237670	-.76924
$-t_{69}t_{53}$		0	0	-.012296	.0048741	0_335228	-.021543	
$-t_{69}t_{48}$		0	0	-.0052726	.0052190	.0021275	-.034309	
$3\,t_{66}t_{59}$		0	-.247237	.027655	-.456357	-.428244	-.045029	
$\bar{s}_6 = t_7 t_{89}$		.42836	.1423	.0990886	.598804	.277157	.00508568	
$\bar{s}_6 = \Sigma$		.42836	-.13301	.106175	.52240	-.49605	-.092805	.02713

divided into two parts, viz.

(i) the intrinsic secondary contribution by the surface;

(ii) the secondary contribution arising from the primary imperfections of the incident pencil of rays.

Similar subdivisions may also be carried out for the tertiary and higher order contributions.

The table has been arranged so as to explain itself, provided reference be made to eqs.(20.41 - 431) and (21.2). A current symbol t_ν is attached to various entries so as to facilitate reference to them when they subsequently recur in the table. The entries are reduced to six significant figures. In the case of corrected systems the number of significant figures carries should exceed by at least two the number of significant figures desired in the final result. Since no trigonometrical or other tables are required anywhere this creates no practical difficulty.

There are about 200 entries per surface consisting entirely of simple additions and multiplications. If a high speed calculating machine is available these can be carried out very rapidly, particularly when it is noticed that certain entries are later multiplied by a variety of other entries, so that the former may be left on the machine for a considerable time. The greater the "memory" of the machine the more can the number of entries be reduced. But in any case a computing time of two to three hours per surface seems reasonable. It may also be noticed that the requisite number of entries is equivalent to that of the tracing of about six skew rays (using Conrady's scheme (v. §2)); but in the latter, trigonometrical tables are extensively used.

The final result of Table I is

$$
\begin{aligned}
\underline{\varepsilon}_k{}' = \;& + 1.3592\,\underline{Y}_1\xi - 0.16\,857\,\underline{V}_1\xi - 0.33713\,\underline{Y}_1\eta + .009739\,\underline{V}_1\eta \\
& + 0.17495\,\underline{Y}_1\zeta - 0.35468\,\underline{V}_1\zeta - 90.921\,\underline{Y}_1\xi^2 - 12.877\,\underline{V}_1\xi^2 \\
& - 50.960\,\underline{Y}_1\xi\eta + 4.3946\,\underline{V}_1\xi\eta - 0.65780\,\underline{Y}_1\xi\zeta + 1.5376\,\underline{V}_1\xi\zeta \\
& + 3.0157\,\underline{Y}_1\eta^2 + 1.7064\,\underline{V}_1\eta^2 + 3.4514\,\underline{Y}_1\eta\zeta - 0.83821\,\underline{V}_1\eta\zeta \\
& - 0.70824\,\underline{Y}_1\zeta^2 + .02713\,\underline{V}_1\zeta^2.
\end{aligned}
\tag{25.2}
$$

(c) TABLE II.

This contains the additional calculations required for the determination of the tertiary spherical aberration. There are about 50 entries per surface. As was explained in T§11.c it is frequently advisable to compute the tertiary spherical aberration in addition to the general secondary aberrations, since the former is the dominant term amongst the set of tertiary terms when the inclination of the principal ray to the axis of the system is sufficiently small. More generally it will be worth while to calculate the spherical aberration of order n+1 when the full aberrations are known up to order n; for this will be a simple matter compared with the determination of the full (n+1)th order terms.

The last row gives the contributions by the surfaces to the tertiary spherical aberration, $t_{1j}{}'$ ($t_{1pj}{}'$ here!); and $T_{1k}{}'$

TABLE II.

	$j \longrightarrow$		1	2	3	4	5	6	Σ_1^k
s_{1q}	$t_2\ t_{22}$		17.7842	13.6036	17.3455	7.61051	0494562	2.92819	
	$t_3\ t_{21}\ t_{26}$		5.47055	-12.8791	-14.2447	6.64979	0349418	-.919811	
	$\frac{1}{2}\Sigma_1^j$	t_{109}	11.6374	.36272	1.63930	7.13015	0406990	.954190	
	$t_5\ t_{16}\ t_{109}$		-9.84311	1.68633	-17.0278	-11.2445	$0_4$404704	-.166182	
	$t_{49}\ t_{73}$		-55.4278	.786281	-.427723	26.6103	$-0_3$177621	-.00105	
	Σ	t_{110}	-65.2704	2.37261	-17.4575	14.3658	$-0_3$137161	-.167236	
	$(t_6)^2$	t_{111}	3.37328	2.44476	3.09446	2.42589	0014493	.717396	
	$t_1\ t_6$	t_{112}	0	2.87599	-5.98614	-2.54031	00675481	.129597	
t_{1p}	$5\ (t_{111})^2 t_9$	t_{113}	524.204	310.977	710.826	342.470	$0_5$157440	7.85772	
	$3\ (t_{111})^2 t_{12}$	t_{114}	0	60.6639	332.659	47.7067	$0_6$442350	0361467	
	$12\ t_{111}(t_9)^2$	t_{115}	6449.72	3176.81	8184.91	3943.43	$0_4$191451	80.2711	
	$10\ t_{111} t_9\ t_{112}$	t_{116}	0	731.661	-2750.14	-722.896	$0_5$485069	2.83893	
	$3\ t_{111}\ t_9\ t_{12}$		0	358.215	1576.02	238.386	$0_3$188280	.153853	
	$(8t_{111}-6t_{12})\ t_{112}\ t_{12}$	t_{117}	0	-7.21606	8100.27	-22.0975	$-0_6$390180	.0160868	
	$3\ t_{111}(t_{12})^2$	t_{118}	0	53.9623	1244.87	53.1402	$0_7$740676	0017962	
	$6\ (t_9)^3$		14932.7	5634.19	16362.2	7883.20	$0_4$339599	142.364	
	$10\ (t_9)^2 t_{12}$	t_{119}	0	8114.21	-13874.6	-3468.29	$0_4$206465	12.0841	
	$-9\ (t_9)^2 t_{12}$		0	-3297.27	-22972.0	-3204.43	$-0_4$240431	-1.96469	
	$-4\ t_9 t_{112} t_{12}$	t_{120}	0	-405.016	416.61	322.092	$-0_5$324832	-.0370591	
	$3\ t_9\ (t_{12})^2$		0	257.342	6972.687	264.956	$0_3$16226	.00502102	
	$(t_{12})^3$		0	38.7269	1552.84	19.7309	$0_9$413288	$0_3$129219	
	Σ		24176.6	10057.3	4257.08	6687.91	$0_4$594292	243.627	
	$\frac{1}{16}\ t_6\ \Sigma$	t_{121}	2779.36	982.580	-468.041	-543.957	$0_6$247754	12.8969	
	$\frac{5}{2}(t_{111})^3$		96.8173	36.5298	74.0787	35.6906	$0_8$220182	.923028	
	$\frac{1}{8}\ t_{113}$		1153.87	435.368	995.156	479.459	$0_3$262416	11.0008	
	$8\ (t_{111})^2 t_{112}$		0	137.614	-458.670	-120.538	$0_6$911660	.633591	
	$2\ t_{114}$		0	121.328	665.317	95.4133	$0_6$884700	0722933	
	$\frac{1}{3}\ t_{115}$		2806.57	1058.94	2758.30	1314.48	$0_5$638270	26.7570	
	$\frac{13}{10}\ t_{116}$		0	951.157	-3575.19	-939.763	$0_5$630577	3.69067	
	$-\frac{1}{2}\ t_{117}$		0	3.60803	-1550.13	11.0488	$0_6$195090	-.00847837	
	$-\frac{4}{3}\ t_{118}$		0	-111.936	-1659.82	-70.8536	$-0_6$987436	-.00157252	
	$\frac{3}{10}\ t_{119}$		0	934.292	-8958.87	-1040.49	$0_5$619395	3.62523	
	$\frac{3}{2}\ t_{120}$		0	-607.523	6174.91	483.137	$-0_5$487248	-.055559	
	Σ			2969.28	-564.314	247.583	$0_4$178683	46.5870	
	$t_{121} - 8t_1\ \Sigma$	t_{122}	2779.36	302.436	-228.000	-594.830	$0_3$550987	12.0068	
S_{1p}	$\Sigma_1^{j-1}(t_{123})$	t_{124}	0	176.654	740.305	-1.1149	-99.3324	-90.3510	
s_{1q}	t_{110}		-65.2704	2.37261	-17.4575	14.3658	$-0_3$137161	-.167236	
	$-3\ t_{61}\ t_{49}$		0	.946541	-1.47385	51.3066	-.0172910	$-0_3$7620	
	$(t_{50}+t_{55})\ t_{58}$		0	-3.13891	16.5080	-13.3935	-.0337289	.283384	
	Σ	t_{125}	-65.2704	.18024	-2.42247	52.2789	-.051151	.115380	
S_{1q}	$\Sigma_1^{j-1}\ t_{125}$	t_{126}	0	-65.2704	-65.0902	-67.5127	-15.2338	-15.2860	
	t_{61}^2		0	19.6862	19.0602	30.0555	.434534	.446487	
	$-\ t_{126}$		0	65.2704	65.0902	67.5127	15.2338	15.2860	
	$3\ \Sigma$	t_{130}	0	254.870	262.451	262.704	47.0060	47.1944	
	$-2\ t_{58}\ t_{61}$	t_{131}	0	117.528	269.622	62.4789	-.670696	-.679441	
	$t_{124}+t_{131}$	t_{132}	0	294.182	1009.93	61.3640	-100.023	-100.030	
t_{1p}	$t_5\ t_{122}$		2779.36	1645.11	-669.674	-497.221	$0_7$310197	18.2550	
	$-5\ t_{61}\ t_{127}$		0	5621.75	-146.693	-1106.34	$0_3$6774	8.1442	
	$(t_{128}+t_{129})\ t_{58}$		0	2465.81	1954.83	871.695	$0_4$41683	-3.94452	
	$t_{46}\ t_{130}$		0	4494.53	-6034.30	-900.14	00505978	88.1409	
	$(t_{47}+t_{52})\ t_{132}$		0	3679.76	-13789.9	454.956	.227644	-146.419	
	$(t_{58})^2(t_{48}+t_{53})$		0	523.083	-2764.59	-164.478	.0194657	.170345	
	Σ		2779.36	16450.4	-21451.4	-2407.49	.252217	-25.6523	-4653.47

itself in the seventh column. Hence we have

$$\underline{\text{sph.abn.}} = +1.3592\, Y_1^3 - 90.921\, Y_1^5 - 4653.5\, Y_1^7 + \underline{O}(9) \,. \qquad (25.3)$$

(We have put $Z_1 = \underline{V}_1 = 0$).

Rejecting $\underline{O}(9)$ we deduce for the positive maximum of the spherical aberration

$$\left.\begin{array}{l} \underline{\text{max.}}\ \underline{\text{sph.}}\ \underline{\text{abn.}} = +.0301 \\ \text{which occurs } \underline{\text{when}} \\ \qquad Y_1 = 7.88 \,. \end{array}\right\} \quad (25.31)$$

(All numerical values are given as % focal length).

As nearly as it is possible to conclude from a considerable number of strict trigonometrical traces the actual maximum is

$$\left.\begin{array}{l} \underline{\text{max.sph.abn.}} = +.0280 \\ \quad \underline{\text{when}}\ Y_1 = 7.58 \,. \end{array}\right\} \quad (25.32)$$

The stop number corresponding to this zone is f'/6.60. These figures show how favourably the information supplied by (25.3) compares with that obtained by trigonometrical tracing. They may also be compared with the results of the usual primary theory (i.e. $\underline{\text{sph.abn.}} = 1.3592\, Y_1^3$) which, of course, gives no maximum at all, and at $Y_1 = .0788$ gives the value $+ .0665$.

(d) FIGURES 1 - 8.

To obtain easily visualised comparison between the results predicted by means of (25.2), (25.3) and (7.31) and those obtained by ordinary trigonometrical tracing one may either plot ε_y' and ε_z' separately as functions of certain parameters at the first surface; or else we may trace a curve which corresponds to the intersection points of a chosen family of rays with the image plane. We shall pursue the latter course below, by considering in Figs.1-6 the family of rays (from a given object point) generated by varying θ for fixed ρ (v. (6.2)); whilst in Figs.7 and 8 the family is generated by varying ρ for fixed θ.

In the various diagrams the meaning of the different curves is as follows:

1.	....	primary aberrations.
2.	----------	(primary + secondary) aberrations
3.	-.-.-.-.-.-	Fig. 1 (primary + secondary + tertiary) aberrations Figs. 2-8 (full primary + secondary) + tertiary spherical aberrations.
4.	________	curve obtained from a number of strict trigonometrical traces.

In the diagrams ε_y' and ε_z' are measured in % focal length, as are all linear measures below. The origin corresponds to the ideal image point, whilst " pr " is the intersection point of the principal ray. All diagrams are drawn to the same scale, and numbers along the curves specify particular rays of the family.

(i) Fig.1

This represents the spherical aberration when the system is working at its maximum aperture ($f'/5.6$, or $\rho = 8.91$). At first sight the convergence of successive approximations towards the "true" curve seems rather slow. But this is explained by the fact that the aperture of the diaphragm is such that the focus of the extreme (i.e. marginal) ray nearly coincides with the paraxial focus. (The vast improvement which occurs if the aperture is made slightly smaller is exemplified by Figs. 2 and 3.)

(ii) Fig. 2.

This again represents a family of extreme rays of a pencil the principal ray of which makes an angle of 6° with the axis of the system in the object space. Here ρ (= 8.27) has its maximum value consistent with the passage of the pencil through the system without vignetting. The corresponding stop-number is $f'/6$. It will be seen that the inclusion of the tertiary spherical aberration results in a further striking improvement in the accuracy of the predicted aberrations.

(iii) Fig. 3.

As Fig. 2, but with an aperture ($\rho = 5.51$) which is two-thirds of that used there; this is equivalent to an effective stop-number $f'/9$. The third curve coincides with the "true" curve. Figs. 2 and 3 illustrate the remarkable deterioration of the image which occurs when the aperture is increased beyond a certain size.

(iv) Fig. 4.

A family of extreme rays of a 12° pencil. As before ρ (= 7.65) has its maximum value corresponding to an effective aperture $f'/6.5$.

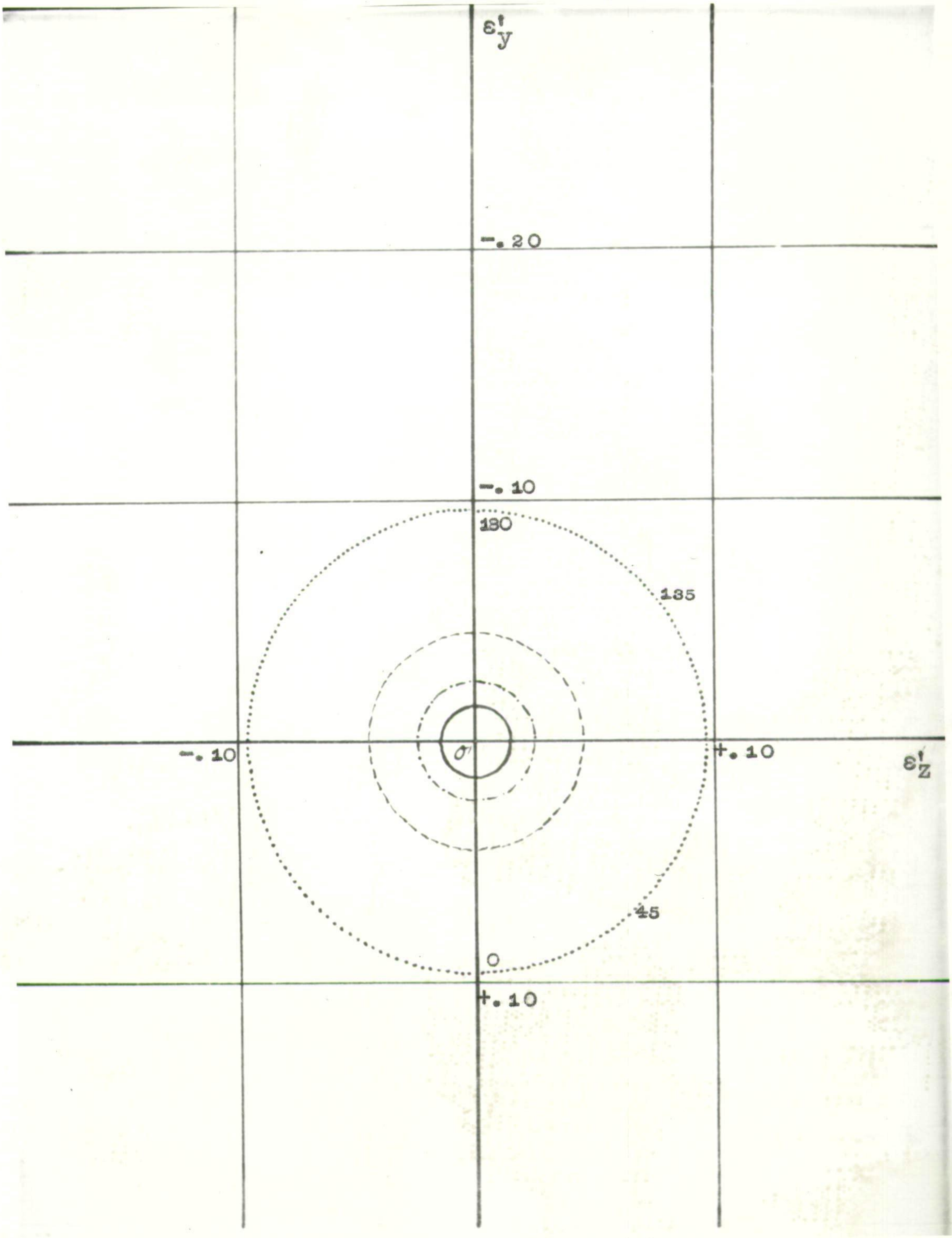


Figure 1. 0^{o} pencil (spherical aberration). $\rho = f'/5.6$.

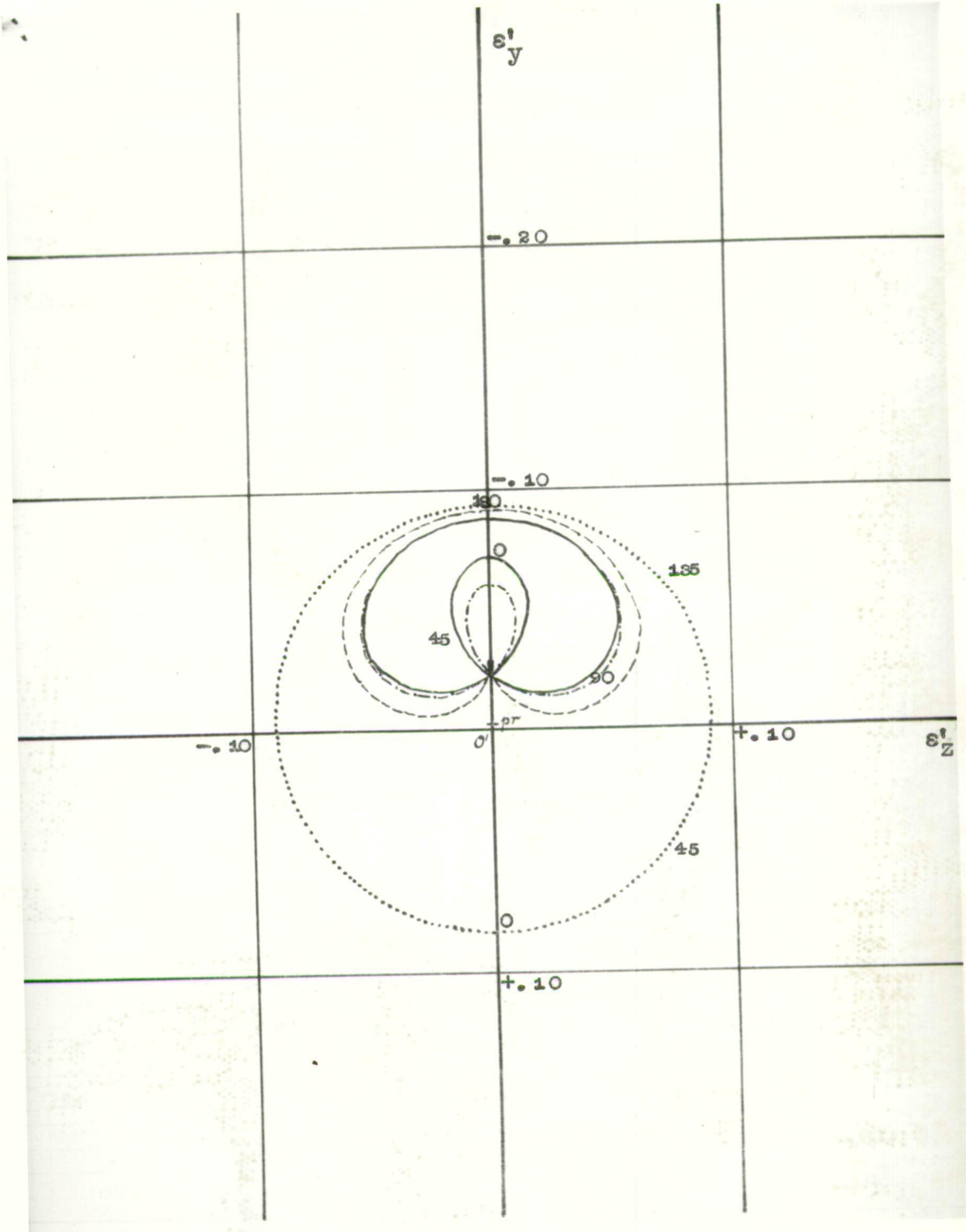


Figure 2. 6° pencil. $\rho = f'/6$.

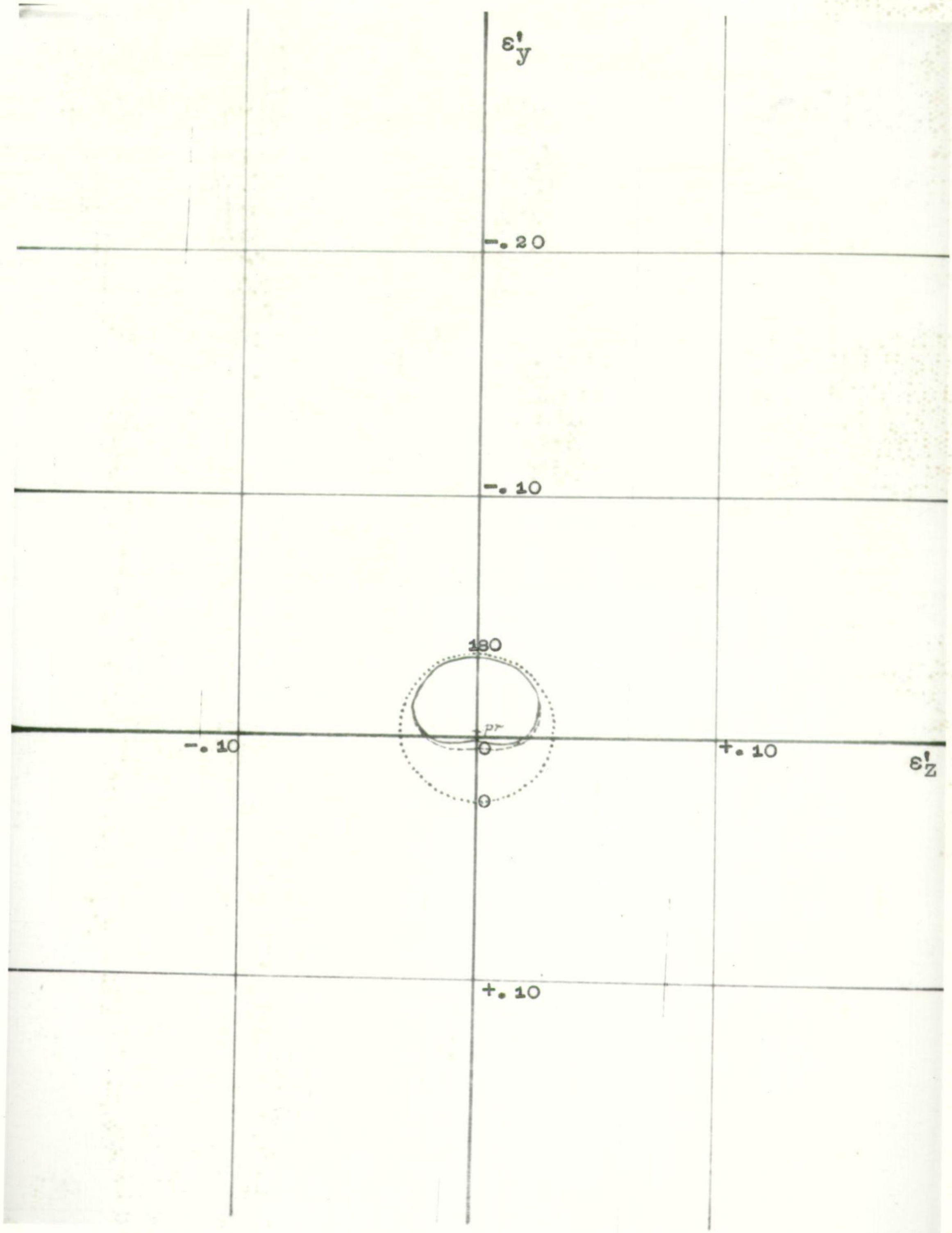


Figure 8. 6° pencil. $\rho = f'/9$.

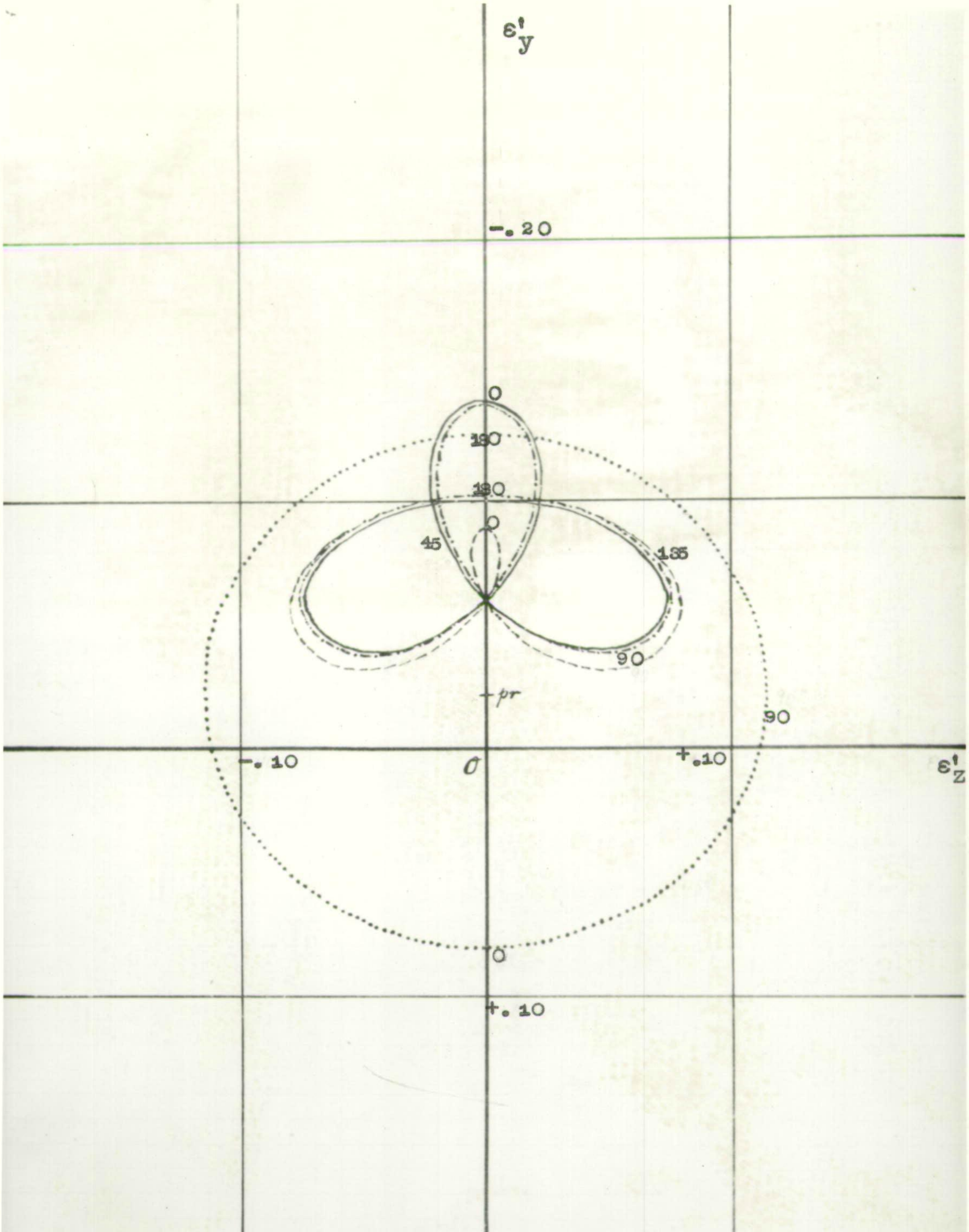


Figure 4. 12° pencil. $\rho = f'/6.5$.

(v) Fig. 5.

As Fig. 4 but the curves are the intersection points of the rays with an out-of-focus plane. The position of the latter is defined by $\underline{x}' = -.516\,9$, i.e. the plane lies nearer to the system than the ideal image plane by a distance $.517\ \%\, f'$. O' is the projection of the ideal image point on to this plane. O_x' is the "new ideal image" point defined by (7.22); it is displaced with respect to O' through a distance

$$O'O_x' \quad = \quad x'V_1\ (u_{qk}' - \rho_q u_{pk}'/\rho_p) \quad = \quad -\ .1134\ (\ \%\, f'\ ) \ , \qquad (25.4)$$

$$\text{since} \quad \rho_p \quad = \quad .8492$$

$$\rho_q \quad = \quad -\ .09615$$

$$\text{and} \quad V_1 \quad = \quad \tan 12^\circ \quad = \quad .2126\ .$$

The curves were obtained by means of the approximate equation (7.31).

(vi) Fig.6.

Extreme rays of an 18° pencil. ρ (= 6.41) is again the maximum aperture, equivalent to an effective stop-number $f'/7.8$. It will be seen that the inclusion of the tertiary spherical aberration makes the predicted aberrations considerably worse than the values given by the full primary and secondary terms alone. But this is of course to be expected. Indeed too much importance should not be attached to the excellent fit between the "third" and the "true" curves in the case of the 12° pencil. However, the fit between the second and the true curves is good in every case; and the single tertiary term does what is demanded of it, i.e. it gives an indication in

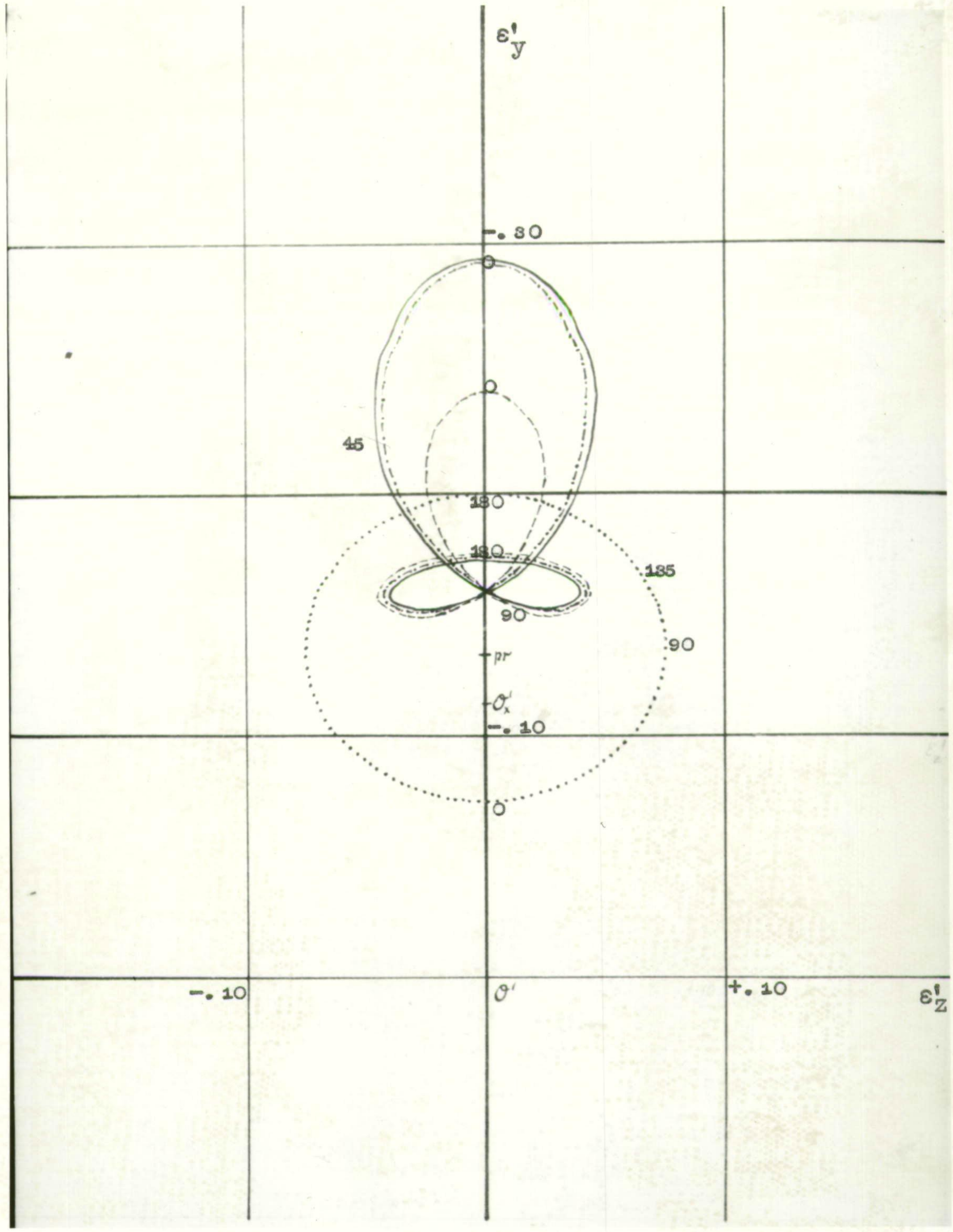


Figure 5. 12^{0} pencil (out-of-focus image)
$\rho = f'/6.5$. ($x' = -0.52f'$)

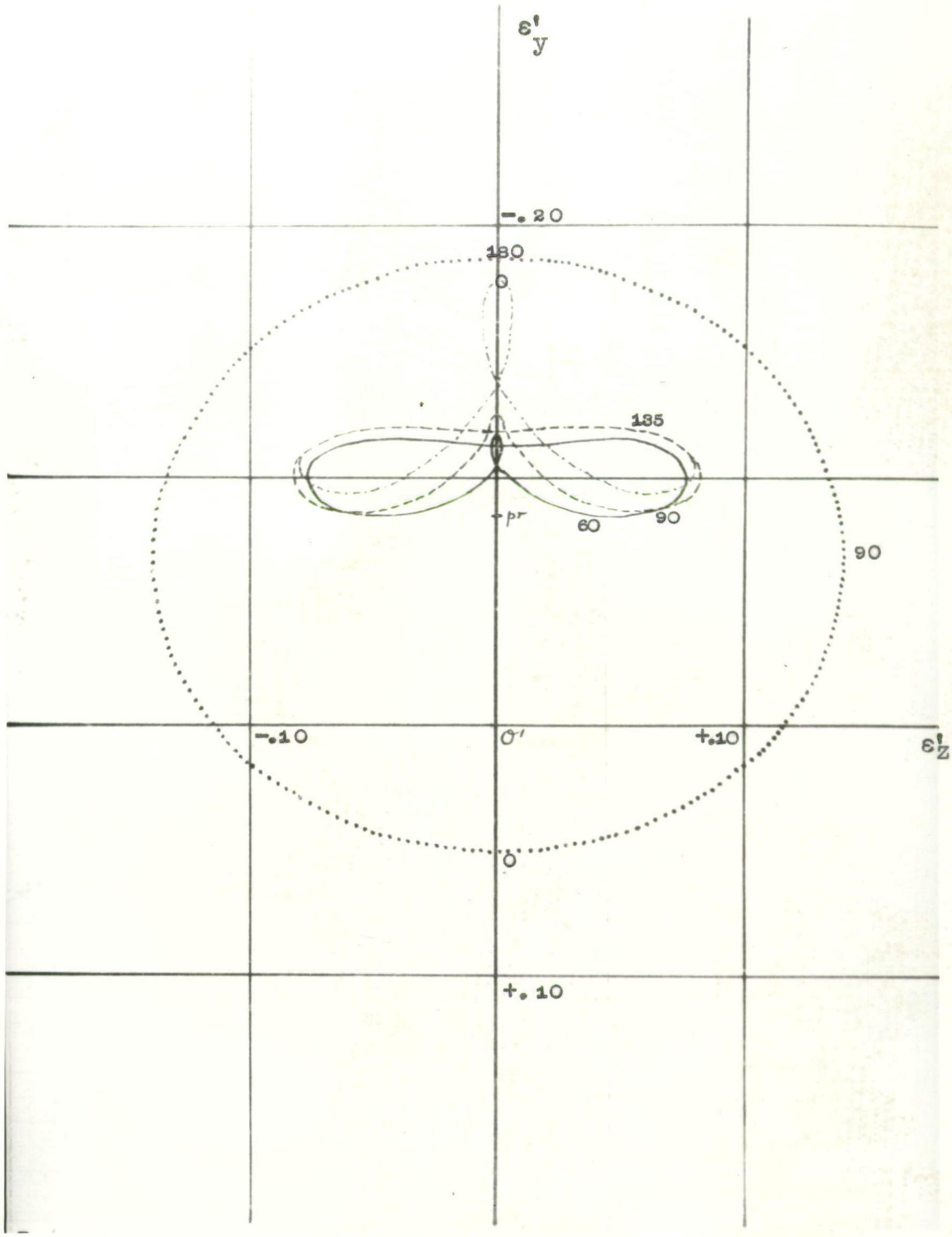


Figure 6. 18° pencil. $\rho = f'/7.8$.

a general way of the difference between the predicted and the true curves for small inclinations of the principal ray.

(vii) Fig. 7.

A family of sagittal rays ($\theta = 90^{\circ}$, ρ variable) of the 12° pencil.

(viii) Fig. 8.

A family of intermediate rays ($\theta = 30^{\circ}$, ρ variable) of the 12° pencil. This diagram strikingly illustrates the point that the inclusion of secondary terms really removes two restrictions implicit in primary theory; viz. (a) the inclination of the principal ray need no longer be infinitesimal; The different curves are reasonably close together when ρ is small enough; implying that a 12° inclination of the principal ray is to be regarded as qualitatively infinitesimal in this optical system. But when the aperture is opened up beyond about $f'/11$ the curves very much part company; and the aperture can no longer be regarded as "infinitesimal" in any sense.

§26. Application of Identities. (Checks).

It is of some interest to compare the numerical values of the two sides of the identities of §17 by using the values of the quantities appearing in Tables I and II. In Table III below the two sides are placed side by side. All coefficients refer to the image space so that the dash and subscript 6 may be omitted.

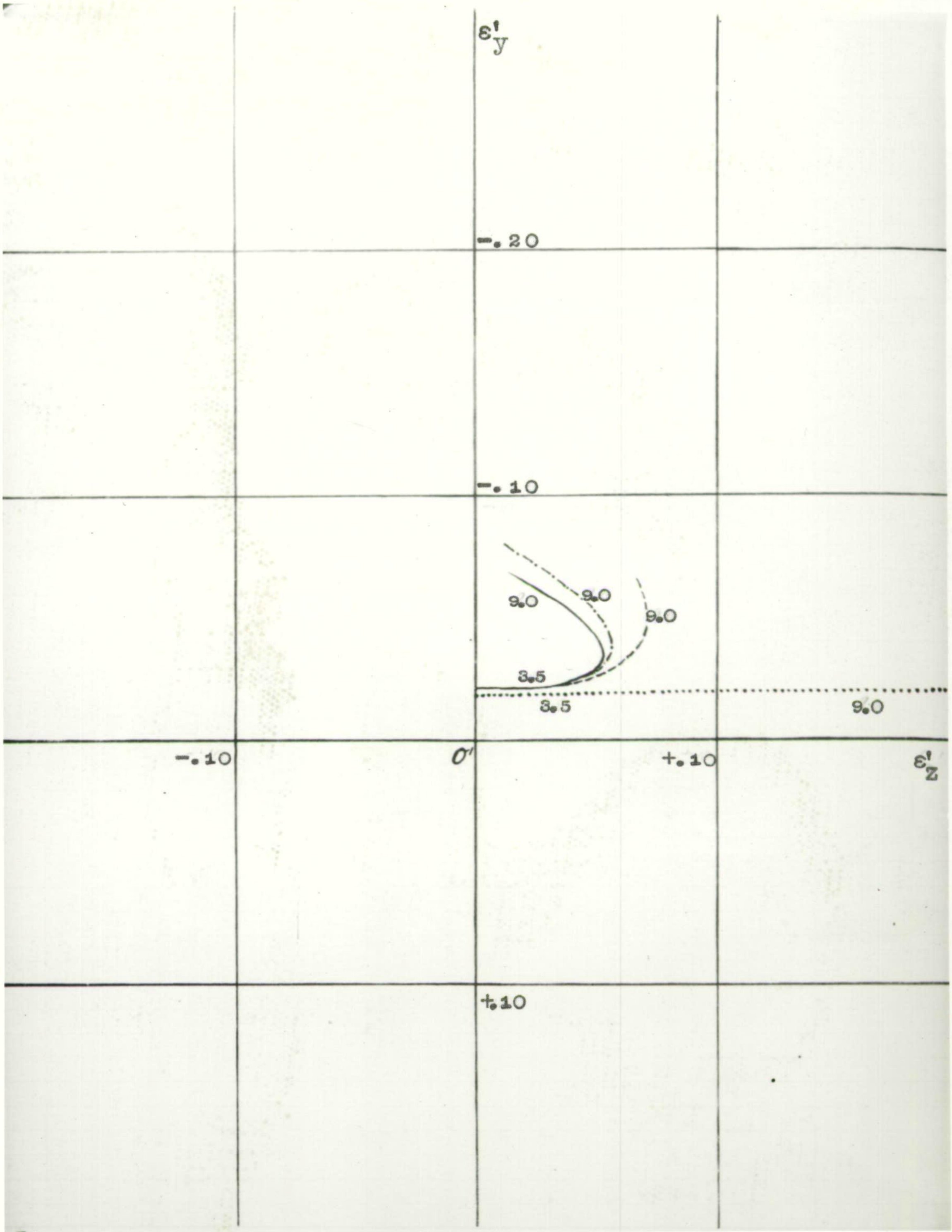


Figure 7. 12° sagittal pencil. $\Theta = 90°$.

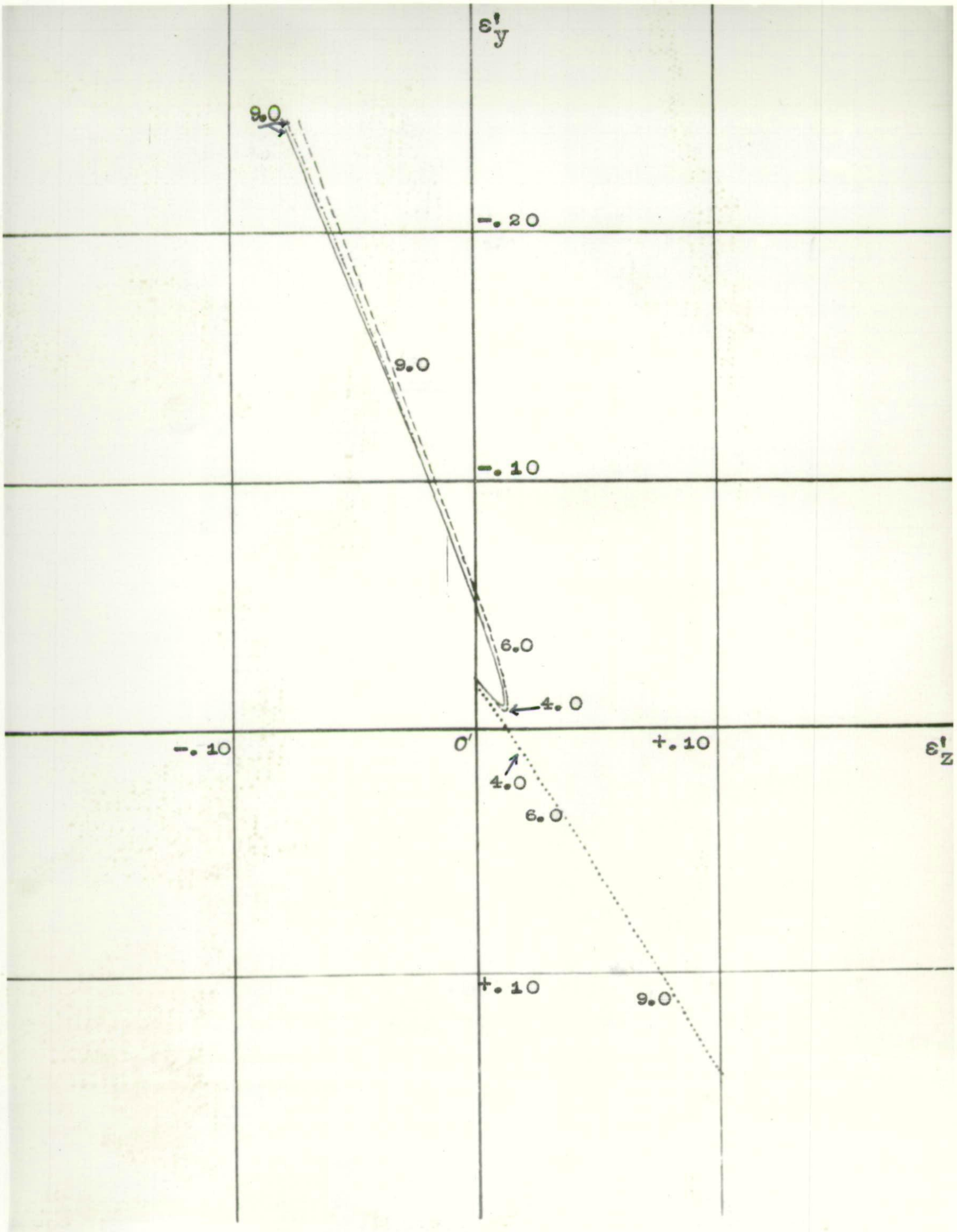


Figure 8. 12° 'intermediate' pencil. $\theta = 30°$.

L. H. S.	Value	R. H. S.	Value
$\overline{A}_p - A_q$	+ .500010	$\frac{1}{2} N_1 u_p^2$	+ .500000
$\overline{B}_p - B_q$	+ .918976	$N_1 u_p u_q$	+ .918978
$\overline{C}_p - C_q$	- .077740	$\frac{1}{2} N_1 (u_q^2 - 1)$	- .077740
$2\overline{A}_p - B_p$	- .000001	0	0
$\overline{B}_q - 2C_q$	+ .000002	0	0
$C_p - \overline{A}_q$	+ .459488	$\frac{1}{2} N_1 u_p u_q$	+ .459489
$4\overline{S}_{1p} - S_{2p}$	- .5494	ω_7	- .5495
$2\overline{S}_{2p} - 2S_{4p}$	+2.7578	ω_{10}	+2.7582
$2\overline{S}_{3p} - S_{5p}$	- .37618	ω_{13}	- .37614
$\overline{S}_{1p} - S_{1q}$	+2.2923	ω_1	+2.2920

TABLE III

§27. Conclusion.

We have shown how an algebraic theory of the aberrations of optical systems may be built up in such a way as to provide a means of calculating in practice the exact aberrations up to any order desired. This has been actually done for the case of a Cooke Triplet for which the full primary and secondary terms and the tertiary spherical aberration have been determined. The computing scheme shows that the amount of labour involved is such as to make the routine determination of the exact higher order terms feasible. Moreover eqs. (20.41-431) and (21.2) between them represent the explicit algebraic form of the secondary aberrations. This appears to invalidate the contention (Hardy and Perrin, 1932 (b)) that " an extension of the algebraic method to include the fifth order [i.e. fifth degree] terms would be well nigh impossible". The use of quasi-invariants and the heuristic step of concentrating attention on the departure of certain expressions from strict linearity (viz. regarding paraxial equations purely as equations between the first terms of the MacLaurin expansions of certain functions) leads to a powerful methodology applicable to theory and practice alike. Thus the theory of the Seidel aberrations can be represented in a particularly concise and unified manner (§16).

It is hoped to develop along these lines a theory of the aberrations of the general symmetrical optical system at a later date.

The following table contains most of the symbols which occur in the text. The numbers in brackets refer to the paragraph in which the symbol first occurs.

Symbol	Meaning of Symbol.
α	direction cosine of ray w. r. to $\underline{x}$-axis = $\cos U_y \cos U_z$. (3)
$a, \bar{a}$	coefficients of $\underline{Y}_1\xi$, $\underline{V}_1\xi$ in expansion of $\Delta\underline{\Lambda}$. (3)
A_j, $\bar{A}_j$	$\sum_{\nu=1}^{j-1} a_\nu$, $\sum_{\nu=1}^{j-1} \bar{a}_\nu$. (3)
$\underline{a}$(subscript)	refers to ray $(\bar{g}_q/g\ ,\ -\bar{g}_p/g)$. (12)
α_n, $\bar{\alpha}_n$	$F^{(n)'}_{00,k}/\mu'$, $\bar{F}^{(n)'}_{00,k}/\mu'$. (13)
β	neg. dirn. cos. of ray w. r. to $\underline{y}$-axis = $\sin U_y \cos U_z$ (3)
b, $\bar{b}$	coefficients of $\underline{Y}_1\eta$, $\underline{V}_1\eta$ in expansion of $\Delta\underline{\Lambda}$. (3)
B_j, $\bar{B}_j$	$\sum_{\nu=1}^{j-1} b_\nu$, $\sum_{\nu=1}^{j-1} \bar{b}_\nu$. (3)
$\underline{b}$(subscript)	refers to ray $(-g_q/g\ ,+\ g_p/g)$. (12)
β_n	$F^{(n)'}_{10,k}/2\mu'$. (13)
β	$q + \rho\cos\theta$. (16)
γ	neg. dirn. cos. of ray w. r. to $\underline{z}$-axis = $\sin U_z$. (3)
c, $\bar{c}$	coefficients of $\underline{Y}_1\zeta$, $\underline{V}_1\zeta$ in expansion of $\Delta\underline{\Lambda}$. (3).
C_j, $\bar{C}_j$	$\sum_{\nu=1}^{j-1} c_\nu$, $\sum_{\nu=1}^{j-1} \bar{c}_\nu$. (3)
c_o	Nri_o. (3)
$\underline{C}$, $\underline{c}$	$Nr \sin \underline{I}$. (6), (19).
$\underline{\Gamma}_{n,j}$	polynomial of degree (2n+1) in expansion of $\Delta\underline{\Lambda}_j$. (5)
γ_i	terms occurring in expressions for $J^{(1)}$ and $J^{(2)}$. (20)
d'	separation of consecutive surfaces. (2)
Δ	difference symbol. (2)
$\underline{D}_j$	$\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_\nu$. (3)
$\underline{\delta}_{yj}$, $\underline{\delta}_{vj}$	$-\frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{q\nu} + \frac{1}{N_1}\sum_{\nu=1}^{j-1} \Delta\underline{\Lambda}_{p\nu}$. (4)
$\delta\underline{A}_j$	$a_{pj}\,\underline{\delta}_{yj} + a_{qj}\,\underline{\delta}_{vj}$. (4)
d	distance of diaphragm from surface just preceding it. (6)
$\underline{\delta}_{sj}$, $\underline{\delta}_{tj}$	increments in case of general linear co-ordinates. (12)
$\underline{\Delta}$	$\Delta(\underline{\beta} + \sin \underline{I}^*)$. (15)

$\underline{\varepsilon}$	$\underline{H} - \underline{h}$. (3)
η	$Y_1V_1 + Z_1W_1$. (3)
ς	$V_1^2 + W_1^2$. (3)
$\underline{e}$	unit vector in direction of ray. (14)
e	i - i'. (15)
E	$k(\sin^2 I_y + \sin^2 I_z)$. (19)
θ	angle specifying ray at first surface. (6)
θ_i	as θ, but in diaphragm plane. (6)
$f, \overline{f}, F, \overline{F}$	coefficients in linear co-ordinates analogous to $g^{(n)}_{\mu\nu}$, ... (12)
F	E/k. (19)
$\underline{F}$	$H - kE$ (19)
f, f'	focal lengths of system. (12)
$g^{(n)}_{\mu\nu}$, $\overline{g}^{(n)}_{\mu\nu}$	$\underline{n}$th order coefficients in expansion of $\Delta\underline{\Lambda}$. (3)
$G^{(n)}_{\mu\nu}$, $\overline{G}^{(n)}_{\mu\nu}$	$\sum_{\alpha=1}^{j-1} g^{(n)}_{\mu\nu\alpha}, \sum_{\alpha=1}^{j-1} \overline{g}^{(n)}_{\mu\nu\alpha}$. (3)
$g_p, g_q, \overline{g}_p, \overline{g}_q$	constants connected with general linear co-ordinates. (12)
g	$(g_p\overline{g}_q - g_q\overline{g}_p)$. (12)
$\hat{g}^{(n)}_{\mu\nu}$	$\underline{n}$th order coefficients in the expansion of $\underline{J}$. (17)
G	$(1-k)(V \sin I_y + W \sin I_z)$. (19)
$\underline{G}$	$H - E/k$. (19)
H	$\underline{y}$- co-ordinate of intersection point of ray with $\underline{x}$, $\underline{y}$ plane. (2)
$\underline{H}$	co-ordinates of intersection point of ray with l_o-plane. (3)
$\underline{h}$	co-ordinates of (ideal) object point. (3)
H	$V^2 + W^2$. (19)
I	angle of incidence. (2)
$\sin \underline{I}$, $\underline{i}$	$\underline{Y}/r - \underline{V}$. (6), (16)
$\underline{J}$	$\Delta\Lambda_y/C_y$ (= $\Delta\Lambda_z/C_z$). (4)
J_n (p; q; x)	Jacobi polynomial. (22)

$\underline{k}$(subscript)	refers to last surface. (3)
K	$N(VZ - WY)$. (14)
k	N/N'. (16)
$\underline{K}^2$	$-(K/Nr)^2$. (19)
κ	$k/(1-k)^2$. (19)
L	$\underline{x}$- co-ordinate of intersection point of ray with $\underline{x}$, $\underline{y}$ plane. (2)
λ	optical invariant of Lagrange type. (3)
$\underline{\Lambda}$	$Nu_o\underline{H}$. (3)
m, m'	paraxial magnification. (3)
μ, μ'	$N_1 u_{o1}$, $N_k' u_{ok}'$. (3)
$M_n(x)$	coefficient of $\underline{t}^n$ in expansion of $(1 + t + \frac{1}{4} x t^2)^{\frac{1}{2}}$. (22)
$\widetilde{M}_n(x)$	coefficient of $\underline{t}^n$ in expansion of $(1 + t)^{\frac{1}{2}}(1 + x t)^{-\frac{1}{2}}$. (22)
N	refractive index. (2)
ξ	$Y_1^2 + Z_1^2$. (3)
$\underline{n}$	unit outward normal to surface at point of incidence. (14)
O	angle between $\underline{z}$-axis and normal to plane of incidence. (2)
$\underline{O}(n)$	terms of degree not less than $\underline{n}$. (2)
$\underline{o}$(subscript)	refers to ray through axial point of object (Y_o excepted) (3)
$\underline{p}$(subscript)	refers to ray (1, 0). (4)
ϖ	$\frac{1}{r} \Delta \frac{1}{N}$. (4)
p_1	paraxial location of centre of entrance pupil. (7)
π, $\bar{\pi}$	constants occurring in definition of $\underline{S}$, $\underline{T}$. (12)
$\underline{q}$(subscript)	refers to ray (0, 1). (4)
q	i/i_o. (16)
r	radius of surface. (2)
ρ_i	radius of aperture of diaphragm. (6)
ρ	radius of equivalent diaphragm. (6)

ρ_o — $y_{oi} - du_{oi}'$. (6)

$\rho, \bar{\rho}$ — see $\pi, \bar{\pi}$. (12)

$s_i, \bar{s}_i$ $(i=1,\ldots,6)$ — second order coefficients in expansion of $\Delta\underline{\Lambda}$. (3)

$S_{ij}, \bar{S}_{ij}$ — $\sum_{\nu=1}^{j-1} s_{i\nu}, \sum_{\nu=1}^{j-1} \bar{s}_{i\nu}$. (3)

$\sigma_{\mu,\alpha}^{(n)}$ — coefficients occurring in the expansion of $\underline{\chi}_{\mu}^{(n)}$. (8)

s — axial separation of two systems. (11)

$\sigma, \bar{\sigma}$ — see $\pi, \bar{\pi}$. (12)

$\underline{S}$ — general linear co-ordinate. (12)

σ — $\Delta N/\Delta N \cos I$. (14)

S — $\frac{1}{2}c_b(i_o + u_o)(i_o - i_o')(i_o' - u_o)$. (16)

σ_i — Seidel Sums. (16)

$\underline{S}$ — $(\frac{1}{\alpha\sigma} - 1)$. (19)

$t_i, \bar{t}_i$ $(i=1,\ldots,10)$ — third order coefficients in expansion of $\Delta\underline{\Lambda}$. (2)

$\tau_{\mu,\alpha}^{(n)}$ — coefficients in the expansion of $\chi_{\mu x}^{(n)}$. (8)

$\tau, \bar{\tau}$ — see $\pi, \bar{\pi}$. (12)

$\underline{T}$ — general linear co-ordinate. (12)

$\underline{T}$ — $(\frac{\alpha}{\alpha'} - 1)$. (16)

t_ν — current symbol in computing scheme. (25)

U — angle between ray and auxiliary axis. (2)

U_y — angle between $\underline{x}$-axis and projection of ray on $\underline{x}, \underline{y}$ plane. (2)

U_z — angle between ray and projection of ray on $\underline{x}, \underline{y}$ plane. (2)

U_c — angle between auxiliary and $\underline{x}$-axes. (2)

V, v — β/α. (3), (16)

v — $\hat{u}_k'$. (11)

W, w — γ/α. (3), (16)

x, x' — shift of object and image planes. (8), (7)

$\underline{\chi}_{\mu}^{(n)}(\theta)$ — coefficient of ρ^{μ} in the expansion of $\underline{\varepsilon}_k'$. (8)

Y	$\underline{y}$- co-ordiante of intersection pt. of ray with tangent plane. (3)
Y_0	value of Y_1 for principal ray. (7)
Z	$\mathbf{z}$- co-ordinate corresponding to Y. (3)
z	$\hat{y}_k - s\hat{u}_k{}'$.
ψ_n	angle between $\underline{n}$th order asymptotes. (13)
$\Omega^{(n)}_{\mu\nu}$	quantity connected with $\underline{n}$th order identities. (17)
ω_i $(i=1,\ldots,15)$	expression containing only primary coefficients. (17)

$'$	dash denotes quantity <u>after</u> refraction. (2)
(α, β)	a paraxial ray defined by $y_1 = \alpha$, $u_1 = \beta$. (3)
$*$	distinguishes semi-canonical variables and coefficients connected with them. (4)
$\sim$	when placed over $\omega, \Omega, \xi, \eta, \zeta$, the latter refer to general linear co-ordinates. (12)
$[AB]$	$\frac{1}{N_1}\begin{vmatrix} A_p & A_q \\ B_p & B_q \end{vmatrix}$. (17)
$[A]$	$\begin{vmatrix} A_p & A_q \\ u_p & u_q \end{vmatrix}$. (17)
(n) (superscript)	indicates $\underline{n}$th order term in the expansion of the quantity to which it is attached. (19)
$[m]$	$\begin{cases} = n/2 & (n \text{ even}) \\ = (n-1)/2 & (n \text{ odd}) \end{cases}$ (22)
$\binom{n}{m}$	binomial coefficient $\equiv n(n-1)\ldots(n-m+1)/m!$ (22)
<u>Clarendon type</u>	implies that the symbol stands for the $\underline{x}$ and $\underline{y}$ "components" of the quantity it denotes. (3)
<u>Gothic type</u>	in the case of the coefficients $\mathfrak{g}^{(n)}_{\mu\nu}$, $\bar{\mathfrak{g}}^{(n)}_{\mu\nu}$, $\hat{\mathfrak{g}}^{(n)}_{\mu\nu}$ implies that in the expansion in which they occur increments have been neglected. (5)

References.

Buchdahl. 1946 Proc. Phys. Soc. 58. p. 545.

Conrady. 1929 (a) Applied Optics and Optical Design. Oxford Univ. Press. p. 256-272, 307-308.

(b) ibid. §68. (c) ibid. p. 334-339. (d) ibid. p. 111.

Hardy and Perrin. 1932 (a). The Principles of Optics. McGraw - Hill. p. 82.

(b) ibid. p. 90.

Herzberger. 1931. (a) Strahlenoptik. Julius Springer. Berlin.

(b) ibid. p. vii and p. 149. (c) ibid. p. 12.

Madelung. 1943. Die Mathematischen Hilfsmittel des Physikers. Dover Pub. N. Y. p. 57.

Steward. 1926. Proc. Camb. Phil. Soc. XXIII No. 9. §§ 5 ff.

Steward. 1928. (a) The Symmetrical Optical System. Camb. Tracts. No. 25.

(b) ibid. p. 30.

Whittaker. 1915. (a) The Theory of Optical Instruments. Camb. Tracts. No. 7.

(b) ibid. p. 44. (c) ibid. p. 47.

HYLTON BOND.

2nd QUALITY

This article has been removed for copyright or proprietary reasons.

It has been published as: Buchdahl, H. A., 1948. Algebraic theory of the primary aberrations of the symmetrical optical system, Journal of the Optical Society of America, 38(1), 14-19

THE ALGEBRAIC THEORY AND CALCULATION OF THE GEOMETRICAL HIGHER ORDER ABERRATIONS OF OPTICAL SYSTEMS.

II. Aberrations of the general symmetrical optical system.

by H.A. Buchdahl, Department of Physics, University of Tasmania.

Abstract:- The author's algebraic methods for the determination of the higher order aberrations of optical systems are extended to cover symmetrical optical systems containing any number of non-spherical surfaces of revolution. The theory is much the same as for the case of spherical refracting surfaces, and the labour involved in practical calculations is little greater than in that case. It is now possible by such methods to adjust or eliminate aberrations through the manipulation of the asphericities of the surfaces, the corresponding analytical problem consisting merely of the solution of simple rational simultaneous algebraic equations.

On the basis of actual computations certain of the surfaces of a 'corrected' Cooke Triplet are made aspherical so as further to improve the performance of the system. The predicted improvement is compared with the information obtained from strict trigonometrical traces, and is found to agree well with the latter.

§ 1. Introduction.

(a) In previous papers the author developed algebraic methods for the determination of the exact monochromatic higher order aberrations of systems of coaxial spherical refracting surfaces, (Buchdahl, 1946,1947). In the present paper these methods are extended so as to make it possible to deal with symmetrical systems containing any number of non-spherical refracting surfaces. We make full use of results already obtained, and consider only those relations which are new, or differ materially from those already known. All the general remarks of S § 1 remain in force unchanged. (The letter S refers to the later of the two papers quoted above). Indeed, the points concerned with the practical usefulness of the present algebraic methods might legitimately be re-emphasised. Moreover, whereas we previously considered solely the analysis of aberrations of *given* systems, we can now also deal easily deal with the change in aberrations consequent upon certain changes being made in the constitution of the system, viz. any alterations in the *asphericities* of the surfaces; and the problem of effecting the best compromise between them (if the requirements as regards the performance of the system be given) may be reduced to the solution of simultaneous rational algebraic equations. The latter will usually be linear; but occasionally the use of quadratic equations, or of equations of even higher degree, may be of advantage. Moreover the expressions upon which actual calculations are based are conveniently divided into two parts, of which the first comprises those terms which would be calculated if the whole surface were a sphere of radius equal to that of the tangent sphere at the pole of the refracting surface. This has great advantages if the optical system arises as the modification of a system originally containing only spherical surfaces, as will frequently be the case in practice.

(b) The general theory underlying the method remains entirely unchanged. In particular the definitions of the aberration coefficients themselves and of the paraxial coefficients of the system, as well as those of general linear coordinates, and of canonical coordinates continue to apply. (Notice that the paraxial coefficients are unaffected by the asphericities of the surfaces). In short, we can say at once that the whole of S Part I may be taken over unchanged, if we everywhere interpret $\underline{r}$ to mean $\underline{r}_0$, where $\underline{r}_0$ is the radius of the tangent sphere at the pole $\underline{A}$ of the surface, i.e. the paraxial radius of curvature; and provided we abandon all equations which involve the coordinates of Conrady explicitly. (They are superfluous anyhow). In the same way, of S Part II we can retain everything following §17(a); and of S Part III §§ 18 and 21. For all the changes which occur in the various expressions are essentially due merely to the fact that the coefficients $\underline{g}^{(n)}_{\mu\nu,j}$, $\underline{\bar{g}}^{(n)}_{\mu\nu,j}$ of S(5.21) now naturally involve the parameters describing the shape of the surface. We therefore begin by considering the latter in the following section.

§2. Specification of surfaces.

(a) In the case of an aspherical surface the normals to the surface erected at the points of incidence P of different rays do not in general intersect the axis of revolution of the surface in a single point C_o but in different points C. The length PC of the normal itself is of course also variable. We write

$$\left.\begin{aligned} AC &= r \\ PC &= \rho \end{aligned}\right\} \qquad (2.1)$$

In the paraxial limit r and ρ both tend to r_o. It is of course assumed that the surface has no singular point at the pole A, and that it possesses a unique tangent plane at all of its points to be considered subsequently. In accordance with these assumptions we may now specify the surface by means of a set of constants Θ_n, $(n=1,2,\ldots)$, defined by

$$x = \sum_{n=1}^{\infty} \Theta_n \chi^n \quad , \qquad (2.2)$$

where $$\chi = y^2 + z^2 \quad ,$$

x, y, z being current coordinates of points on the surface, in terms of the coordinate system employed previously (v, S§2(a)). The expansion (2.2) allows of the representation of any surface which can occur in an optical system, for x is naturally a single-valued function of χ, but not necessarily vice versa, (e.g. a Schmidt plate). The coefficients Θ_n are unrestricted except insofar as the series (2.2) must converge, (and in practice sufficiently rapidly so for relevant values of χ).

The equation of a sphere of radius r_o is

$$x = r_o(1-\sqrt{1-\chi/r_o^2}) = \frac{1}{2r_o}\chi + \frac{1}{8r_o^3}\chi^2 + \ldots \qquad (2.3)$$

that is, $\theta_n = (-1)^{n-1}\binom{\frac{1}{2}}{n} (2\theta_1)^{2n-1}$, $(n=2,3,\ldots)$. (2.31)

Comparing this with (2.2) we see that the paraxial radius of curvature of the surface is $1/2\theta_1$.

Conicoids of revolution are merely special cases of (2.2). In particular if $\theta_n = 0$, $(n \geqslant 2)$, the surface is a paraboloid of revolution, whilst in the extended paraxial region we may replace any surface for which θ_1 and θ_2 are not zero by its " tangent spheroid"

$$\left.\begin{aligned} &\frac{(x-a)^2}{a^2} + \frac{\chi}{b^2} = 1 , \\ &\text{with} \quad a = \theta_1^2/2\theta_2 \quad \text{and} \quad b^2 = \theta_1/4\theta_2 . \end{aligned}\right\} \quad (2.4)$$

If $\theta_1 = 0$, whilst some other θ_n does not vanish we speak of a <u>figured plane</u>, whilst we call an aspherical surface a <u>figured sphere</u> if $\theta_2 = \theta_1^3 \neq 0$.

(b) In practice a surface will often be specified in a manner differing from (2.2), but that is of course of no consequence. Even here it is convenient to introduce an alternative manner of describing the surface, <u>viz</u>. in terms of the curvature of the curve of intersection of the surface with a plane containing the $\underline{x}$-axis and passing through $\underline{P}$. Denoting the curvature by $\underline{C}$, we write

$$\underline{C} = \sum_{n=0}^{\infty}(2n+1)a_n\chi^n \quad , \qquad (2.5)$$

where the factor $(2n+1)$ is inserted for later convenience. The sign of $\underline{C}$ is fixed by the convention

$$a_0 = 1/r_0 . \qquad (2.51)$$

In terms of an auxiliary variable $\underline{t}$, the coefficients θ_n and a_n are then related by the equation

$$\sum_{s=0}^{\infty}(2s+1)\alpha_s t^s = \frac{d}{dt}\frac{T}{\sqrt{1+T^2}} \quad , \tag{2.6}$$

where $$T = \sum_{s=1}^{\infty} 2s\theta_s t^{2s-1} \quad . \tag{2.61}$$

In particular we have

$$\left.\begin{aligned} \alpha_0 &= 2\theta_1 \\ \alpha_1 &= 4(\theta_2 - \theta_1^3) \\ \alpha_2 &= 6(\theta_3 - 4\theta_1^2\theta_2 + 2\theta_1^5) \\ \alpha_3 &= 4(2\theta_4 - 9\theta_1^2\theta_3 + 30\theta_1^4\theta_2 - 12\theta_1\theta_2^2 - 10\theta_1^7) \quad . \end{aligned}\right\} \tag{2.7}$$

Thus the condition that a surface be a figured sphere is now simply $$\alpha_1 = 0 \quad ; \tag{2.71}$$

whilst for a paraboloid, for instance,

$$\alpha_s = (-1)^s \binom{s-\frac{1}{2}}{s} \alpha_0^{2s+1} \quad . \tag{2.8}$$

The sphere itself is, of course, characterised by

$$\alpha_s = 0 \quad , (s>0) \; . \tag{2.9}$$

§3. Refraction at a surface.

(a) In this section, which corresponds to S§14, we consider the equations which govern the refraction of rays at a given surface. These then lead to the required expansions in terms of canonical coordinates.

Hereafter we do not in general define again any symbol the meaning of which is the same as that given in the table of S§27. But a table of new symbols is appended in §14.

In considering the unit normal $\underline{n}$ to the surface the complications arising from the asphericity of the surface first appear. Thus now

$$\begin{aligned} \underline{n} &= \left((r-x)/\rho,\ -y/\rho,\ -z/\rho\right) \\ &= \left((r-x)/\rho,\ (xV-Y)/\rho,\ (xW-Z)/\rho\right) \quad . \end{aligned} \tag{3.1}$$

If $\underline{i}$ be the unit normal to the plane of incidence, we have

$$\underline{i} = (\underline{n}\times\underline{e})\operatorname{cosec}I = (\underline{n}\times\underline{e}')\operatorname{cosec}I' = -(\underline{e}'\times\underline{e})\operatorname{cosec}(\Delta I). \quad (3.2)$$

If we define $\tilde{\underline{C}} = N(\underline{Y} - r\underline{V})$,

and, as before, $K = N(WY - VZ)$, (3.3)

then, since $\Delta N\sin I = 0$, and in virtue of (3.1), (3.2) gives rise to the following equations

$$\begin{aligned} \underline{i} &= \frac{\alpha}{N\rho\sin I}(K, -\tilde{C}_z, +\tilde{C}_y) \\ &= \frac{\alpha'}{N'\rho\sin I'}(K', -\tilde{C}'_z, +\tilde{C}'_y) \\ &= \frac{\sigma}{(1-k)\sin I}\Big((\beta'\gamma-\beta\gamma'),-(\alpha\gamma'-\alpha'\gamma),(\alpha\beta'-\alpha'\beta)\Big). \end{aligned} \quad (3.31)$$

This leads at once to

$$\begin{aligned} \Delta\alpha\tilde{\underline{C}} &= 0 \\ \Delta\alpha K = \Delta K^* &= 0 , \end{aligned} \quad (3.4)$$

and
$$\Delta\underline{V} = -\frac{r_0\varpi}{\rho\sigma\alpha'}\tilde{\underline{C}} . \quad (3.5)$$

Since $\underline{i}\cdot\underline{i} = 1$ we have from (3.31)

$$(N\rho\sin I)^2 = \alpha^2(K^2 + \tilde{C}_y^2 + \tilde{C}_z^2) \quad (3.6)$$

Notice that ρ and $\underline{r}$ are connected by the relation

$$(r - x)^2 + \varkappa = \rho^2 . \quad (3.7)$$

(b) From (3.2) and (3.4) it follows that K^* is an optical invariant also in the general symmetrical system, so that S§17(b) remains valid. On the other hand we cannot write down an equation for $\tilde{\underline{C}}$ of the form S(4.53) because $\tilde{\underline{C}}$ is formed from $\underline{Y}$ and $\underline{V}$ with the aid of $\underline{r}$, not $\underline{r}_0$. We therefore define

$$\underline{C} = N(\underline{Y} - r_0\underline{V}) = c_p(\underline{Y}_1 + \underline{\delta}_y) + c_q(\underline{V}_1 + \underline{\delta}_v) . \quad (3.8)$$

At any surface $\alpha\underline{C}$ is however now not invariant; on the contrary

$$\Delta\alpha\underline{C} = (r - r_0)\Delta N\underline{\beta} . \quad (3.81)$$

(c) If we introduce the notation

$$\left.\begin{aligned} \underline{S} &= r/\rho\sigma\alpha' \\ \underline{R} &= 1/r - 1/r_o \quad , \end{aligned}\right\} \quad (3.9)$$

(3.5) becomes $\Delta\underline{V} = (1-k)\underline{S}(\underline{I} + \underline{R}\underline{Y})$, (3.10)

where $\underline{I} = \underline{Y}/r_o - V = C/Nr_o$. (3.101)

Note that $\underline{I}$ now stands for the quantity previously denoted by sin $\underline{I}$, ($\underline{v}$. $\underline{S}$§4(b)).

§4. Expression for $\Delta\Lambda$.

In the detailed treatment of this and the following sections we confine ourselves throughout to the use of canonical coordinates. Nothing stands in the way, of course, of dealing with other linear coordinates in an analogous manner. It will be found that a considerable number of different symbols is introduced below which do not appear in the final expressions. The justification for this procedure lies in its ability to create order out of chaos: it is the simplicity of the final expressions which counts in practical work.

We have $\Delta\underline{\Lambda} = \Delta N(y_o\underline{V} - u_o\underline{Y})$. (4.1)

$$\left.\begin{aligned} &\text{Writing} \quad \Delta N = \overline{N} \ , \\ &\text{and since} \quad \Delta\underline{Y} = x\Delta\underline{V} \ , \ \Delta Nu = \overline{N}y_o/r_o \quad , \end{aligned}\right\} \quad (4.11)$$

$$\begin{aligned} \Delta\underline{\Lambda} &= y_o(\overline{N}\underline{V} + N'\Delta\underline{V}) - (\underline{Y}\overline{N}y_o/r_o + N'u_o'x\Delta\underline{V}) \\ &= -\overline{N}y_o\underline{I} + \overline{N}\underline{S}[\underline{I} + \underline{R}\underline{Y}(y_o - xu_o')] \end{aligned}$$

$$\Delta\underline{\Lambda} = \underline{J}\underline{I} + \underline{L}\underline{Y} \quad , \qquad (4.2)$$

where $\underline{J} = \bar{N}[(\underline{S} - 1)y_0 - \underline{X}u'_0]$, (4.3)

$\underline{L} = \bar{N}R(\underline{S}y_0 - \underline{X}u'_0)$, (4.4)

with $\underline{X} = x\underline{S}$. (4.41)

Notice that $\underline{L} \equiv \underline{R}(\underline{J} + \bar{N}y_0)$. (4.5)

The expression (4.2) shows clearly the quasi-invariant nature of $\Delta\Lambda$, since $(\underline{S} - 1)$, $\underline{X}$ and $\underline{R}$ are all $\underline{O}(2)$. Notice that the second term of $\underline{LY}$ is $\underline{O}(5)$.

§5. Determination of $\underline{S}$.

The only quantity in (4.2) which cannot be directly expanded in terms of the (pre-refraction) canonical variables at the surface, i.e. $\underline{Y}$ and $\underline{V}$, is $\underline{S}$. In this section we therefore derive a quadratic equation for $\underline{S}$, the coefficients of which do not involve $\underline{Y}'$ and $\underline{V}'$; at the same time we avoid the necessity for the explicit determination of α' and of σ.

The equations (3.2) are summarized by

$$\Delta N\underline{e} = \bar{N}\underline{n}/\sigma \quad . \tag{5.1}$$

The $\underline{x}$ component of this gives

$$\Delta N\alpha = \bar{N}(r - x)/\rho\sigma \quad , \tag{5.2}$$

which, in virtue of (3.9) may be written

$$\alpha/\alpha' \equiv \underline{A} = k^{-1}[1 - (1-k)\underline{PS}] \quad , \tag{5.3}$$

where $\underline{P} = (r - x)/r$. (5.31)

Now $\Delta(1/\alpha^2) \equiv \Delta(V^2 + W^2) = (\underline{A}^2 - 1)/\alpha^2$. (5.32)

We have $\Delta(V^2+W^2) = 2(V\Delta V + W\Delta W) + [(\Delta V)^2 + (\Delta W)^2]$. (5.33)

The $\Delta\underline{V}$ may be re-expressed by means of (3.10). To simplify the appearance of the resulting expressions we consider $\underline{I}$, $\underline{Y}$, $\underline{V}$,...

as two-vectors, and write for the scalar product of $\underline{Y}$ and $\underline{V}$ for instance

$$\underline{Y}\cdot\underline{V} = YV + ZW \,. \tag{5.4}$$

(In this notation K is the vector product of $\underline{Y}$ and $N\underline{V}$).
Substituting in (5.32) from (3.10), and eliminating $\underline{A}$ from (5.32) by means of (5.3) we obtain the required quadratic equation for $\underline{S}$, viz.

$$\underline{S}_2\underline{S}^2 + 2\underline{S}_1\underline{S} + \underline{S}_0 = 0 \quad , \tag{5.5}$$

where

$$\left.\begin{aligned} \underline{S}_0 &= -(1-k^2)/k^2a^2 \\ \underline{S}_1 &= (1-k)\left[(\underline{I}+\underline{R}\underline{Y})\cdot\underline{V} + \underline{P}/k^2a^2\right] \\ \underline{S}_2 &= (1-k)^2\left[(\underline{I}+\underline{R}\underline{Y})\cdot(\underline{I}+\underline{R}\underline{Y}) - \underline{P}^2/k^2a^2\right] \,. \end{aligned}\right\} \tag{5.6}$$

§6. Expansions in terms of χ, and of ξ, η, ζ.

(a) Certain functions (e.g. $\underline{R}$) depend upon the point of incidence alone, since they are essentially contained already in the geometry of the surface. Such functions are conveniently first expanded in series of ascending powers of χ by means of the basic expansion (2.2), i.e.

$$x = \theta_1\chi + \theta_2\chi^2 + \theta_3\chi^3 + \underline{O}(8) \,. \tag{6.1}$$

Hereafter we shall, in general, consider only those terms explicitly which suffice for the determination of the exact (primary and) secondary aberrations; moreover we shall use the constants a_n in place of θ_n whenever convenient. Thus, directly,

$$\frac{x}{r} = \frac{d(x^2)/d\chi}{1 + d(x^2)/d\chi} \,. \tag{6.2}$$

Substituting the series (6.1) for x, this gives

$$\underline{P} = 1 - 2\theta_1^2\chi - (6\theta_1\theta_2 - 4\theta_1^4)\chi^2 + O(6) \,. \tag{6.3}$$

On the other hand, multiplying $(1 - \underline{P})$ by $1/x$, we have

$$\underline{R} = a_1\chi + (a_2 + \theta_1^2 a_1)\chi^2 + O(6) . \qquad (6.3)$$

(b) To connect the preceding expansions with the coordinates of the ray before refraction at the surface we express $\varkappa$ in terms of ξ, η, ζ, where

$$\xi = \underline{Y}\cdot\underline{Y} , \quad \eta = \underline{Y}\cdot\underline{V} , \quad \zeta = \underline{V}\cdot\underline{V} . \qquad (6.4)$$

Now from (3.1), since $\underline{n}\cdot\underline{n} = 1$, we get

$$(r-x)^2 + \zeta x^2 - 2\eta x + \xi = \rho^2 ,$$

or, in virtue of (3.7)

$$\zeta x^2 - 2\eta x + (\xi - \varkappa) \equiv 0 . \qquad (6.5)$$

Substituting (6.1) in this identity we find

$$\varkappa = \xi - 2\theta_1\xi\eta + (-2\theta_2\xi^2\eta + \theta_1^2\xi^2\zeta + 4\theta_1^2\xi\eta^2) + \underline{O}(8). \qquad (6.6)$$

With the aid of (6.6) we can now express various quantities which appear in previous sections as series in ascending powers of ξ, η, ζ. Thus

$$x = \theta_1\xi + (\theta_2\xi^2 - 2\theta_1^2\xi\eta) + O(6) . \qquad (6.7)$$

$$\underline{R} = a_1\xi + [(a_2 + \theta_1^2 a_1)\xi^2 - a_0 a_1\xi\eta] + \underline{O}(6) . \qquad (6.8)$$

$$\underline{P} = 1 - 2\theta_1^2\xi + [-(3\theta_1 a_1 + 2\theta_1^4)\xi^2 + 4\theta_1^3\xi\eta] + O(6) . \qquad (6.9)$$

§7. Solution of the equation for $\underline{S}$.

(a) The series of §6(b) may now be substituted in the expressions for $\underline{s}_0$, $\underline{s}_1$, $\underline{s}_2$. This leads then to series in which the variables ξ, η, ζ alone appear, if we remember that

$$\underline{I} = a_0\underline{Y} - \underline{V} . \qquad (7.1)$$

At this stage it is most convenient to introduce the following notation. Let $\underline{B}$ stand for any expression whatever which can be expanded in a series of ascending powers of ξ, η, ζ. Then we write

$$\underline{B} = \sum_{n=0}^{\infty} \underline{B}^{(n)} , \qquad (7.2)$$

where $\underline{B}^{(n)}$ is a homogeneous polynomial of degree $\underline{n}$ in ξ, η, ζ. $\underline{B}^{(n)}$ is therefore $\underline{O}(2n)$ when looked upon as a function of the canonical coordinates. With this notation we find from (5.6) after some straightforward calculation that

$$\left.\begin{aligned}
\underline{s}_0^{(0)} &= -(1-k^2)/k^2 \\
\underline{s}_1^{(0)} &= (1-k)/k^2 \\
\underline{s}_2^{(0)} &= -(1-k)^2/k^2 \\
\underline{s}_0^{(1)} &= \underline{s}_0^{(0)}\zeta \\
\underline{s}_1^{(1)} &= \underline{s}_1^{(0)}[-\tfrac{1}{2}\alpha_0^2\xi + k^2\alpha_0\eta + (1-k^2)\zeta] \\
\underline{s}_2^{(1)} &= -\underline{s}_2^{(0)}[(k^2+1)\alpha_0^2\xi - 2k^2\alpha_0\eta + (k^2-1)\zeta] \\
\underline{s}_0^{(2)} &= 0 \\
\underline{s}_1^{(2)} &= \underline{s}_1^{(0)}[-(\tfrac{3}{2}\theta_1\alpha_1+2\theta_1^4)\xi^2 + (k^2\alpha_1+4\theta_1^3)\xi\eta - 2\theta_1^2\xi\zeta] \\
\underline{s}_2^{(2)} &= -\underline{s}_2^{(0)}[\theta_1\alpha_1(4k^2+3)\xi^2 - (2k^2\alpha_1 + \alpha_0^3)\xi\eta + \alpha_0^2\xi\zeta] .
\end{aligned}\right\} \quad (7.3)$$

(b) Equation (5.5) may now be solved for $\underline{S}$. Expanding $\underline{S}$ according to (7.2) we obtain, since $\underline{S}^{(0)} = 1$,

$$\left.\begin{aligned}
\underline{S}^{(1)} &= -\frac{\underline{s}_0^{(1)} + 2\underline{s}_1^{(1)} + \underline{s}_2^{(1)}}{2(\underline{s}_1^{(0)} + \underline{s}_2^{(0)})} \\
\underline{S}^{(2)} &= -\frac{2\underline{s}_1^{(2)} + \underline{s}_2^{(2)} + 2(\underline{s}_1^{(1)} + \underline{s}_2^{(1)})\underline{S}^{(1)} + \underline{s}_2^{(0)}(\underline{S}^{(1)})^2}{2(\underline{s}_1^{(0)} + \underline{s}_2^{(0)})} .
\end{aligned}\right\} \quad (7.4)$$

Substitution of (7.3) in (7.4) then gives

$$\left.\begin{aligned}
\underline{S}^{(1)} &= \tfrac{1}{2}[(k^2-k+1)\alpha_0^2\xi - 2k^2\alpha_0\eta + k(k+1)\zeta] \\
\underline{S}^{(2)} &= [\tfrac{1}{4}(4k^2-4k+3)\alpha_0\alpha_1 + 2(3k^4-5k^3+7k^2-5k+3)\theta_1^4]\xi^2 \\
&\quad -[k^2\alpha_1 + \tfrac{1}{2}(3k^4-3k^3+3k^2-k+1)\alpha_0^3]\xi\eta + k(3k^3-k^2+2)\theta_1^2\xi\zeta \\
&\quad + 2k^3(3k-1)\theta_1^2\eta^2 - k^2(k+1)(3k-2)\theta_1\eta\zeta - \tfrac{3}{8}k^2(k+1)(1-k^2)\zeta^2 .
\end{aligned}\right\} \quad (7.5)$$

Accordingly the expansion of $\underline{X}$ is, by (4.41) and (6.7),

$$\left.\begin{aligned} \underline{X}^{(1)} &= \tfrac{1}{2}a_0\xi \\ \underline{X}^{(2)} &= \tfrac{1}{2}a_0\xi\underline{S}^{(1)} + \left[(\tfrac{1}{4}a_1 + \tfrac{1}{8}a_0^3)\xi^2 - \tfrac{1}{2}a_0^2\xi\eta\right] \ . \end{aligned}\right\} \quad (7.6)$$

It will be noticed that the main burden of deriving the expressions for the calculation of aberrations up to order $\underline{n}$ lies in the determination of $\underline{S}^{(1)},\ldots,\underline{S}^{(n)}$.

§8. Introduction of the Υ_i.

In $\underline{S}$§20(b) certain expressions $\Upsilon_1,\ldots,\Upsilon_4$ were introduced which made it possible to express $\underline{\Delta\Lambda}$ in a compact form, especially suitable for the purposes of actual computation. We introduce them again here, but this time directly in terms of ξ, η, ζ. We also introduce three further symbols $\Upsilon_5, \Upsilon_6, \Upsilon_7$ to express those parts of $\underline{\Delta\Lambda}$ which vanish when the surface is spherical. Thus

$$\left.\begin{aligned} \Upsilon_1 &= \tfrac{1}{2}(k^2-k+1)a_0^2\xi - k^2a_0\eta + \tfrac{1}{2}k(k+1)\zeta \\ \Upsilon_2 &= -\tfrac{1}{2}(k-1)^2a_0^2\xi + k(k+1)a_0\eta + \tfrac{1}{2}(1-k^2)\zeta \\ \Upsilon_3 &= \tfrac{1}{4}(2k^2-k+2)a_0^2\xi - \tfrac{1}{2}k(2k+1)a_0\eta + \tfrac{1}{4}(2k^2+3k-3)\zeta \\ \Upsilon_4 &= -\tfrac{1}{4}(k-1)^2a_0^2\xi + \tfrac{1}{2}(k-1)(k-2)a_0\eta - \tfrac{1}{4}(1-k)(3-k)\zeta \\ \Upsilon_5 &= \xi \\ \Upsilon_6 &= (k^2-k+\tfrac{3}{4})a_0\xi - k^2\eta \\ \Upsilon_7 &= \Upsilon_1 + a_0^2\xi - a_0\eta \ . \end{aligned}\right\} \quad (8.1)$$

We then have

$$\left.\begin{aligned} \underline{S}^{(1)} &= \Upsilon_1 \\ \underline{S}^{(2)} &= \Upsilon_1(\Upsilon_3 - \Upsilon_4) + \kappa\Upsilon_2\Upsilon_4 + a_1\Upsilon_5\Upsilon_6 \quad , \; (\kappa = k/(1-k)^2 \;), \end{aligned}\right\}(8.2)$$

$$\left.\begin{aligned} \underline{X}^{(1)} &= \tfrac{1}{2}a_0\Upsilon_5 \\ \underline{X}^{(2)} &= \tfrac{1}{2}a_0\Upsilon_5\Upsilon_7 + \tfrac{1}{4}(a_1 - \tfrac{3}{2}a_0^3)\Upsilon_5^2 \end{aligned}\right\}(8.3)$$

$$\left.\begin{aligned} \underline{R}^{(1)} &= a_1\Upsilon_5 \\ \underline{R}^{(2)} &= (a_2 - \tfrac{3}{4}a_0^2a_1)\Upsilon_5^2 + a_1(\Upsilon_7 - \Upsilon_1)\Upsilon_5 \;. \end{aligned}\right\}(8.4)$$

§9. Final expressions for $\Delta\Lambda$.

(a) The results of the previous section may now be applied to (4.3) and (4.4). We have

$$\left.\begin{aligned} \underline{J}^{(1)} &= \bar{N}\,(\underline{S}^{(1)}y_0 - \underline{X}^{(1)}u_0') \\ \underline{J}^{(2)} &= \bar{N}\,(\underline{S}^{(2)}y_0 - \underline{X}^{(2)}u_0') \end{aligned}\right\}(9.1)$$

$$\left.\begin{aligned} \underline{L}^{(1)} &= \bar{N}\,\underline{R}^{(1)}y_0 \\ \underline{L}^{(2)} &= \bar{N}[(\underline{R}^{(2)} + \underline{R}^{(1)}\underline{S}^{(1)})y_0 - \underline{R}^{(1)}\underline{X}^{(1)}u_0'] \;. \end{aligned}\right\}(9.2)$$

It is obvious that when $a_1 = 0$ the expressions for $\underline{J}^{(1)}$ and $\underline{J}^{(2)}$ must reduce to those given by $\underline{S}$(20.41) and $\underline{S}$(20.42), if the latter are first multiplied by Nr_0. This may be confirmed to be the case. Hence we write

$$\left.\begin{aligned} \underline{J}_S^{(1)} &= Nr_0[e_0\Upsilon_1 + u_0'\Upsilon_2] \quad , \\ \underline{J}_S^{(2)} &= Nr_0[\Upsilon_3\underline{J}_S^{(1)} + \Upsilon_4(u_0\Upsilon_1 + \kappa e_0\Upsilon_2)] \;. \end{aligned}\right\}(9.3)$$

Such additional terms as appear in $\underline{J}^{(2)}$ and all the terms of $\underline{L}^{(1)}$ and $\underline{L}^{(2)}$ must then contain a_1, a_2 as factors. In fact we find

$$\left.\begin{aligned} \underline{J}^{(1)} &= \underline{J}_S^{(1)} \\ \underline{L}^{(1)} &= \bar{a}_1\,y_0\Upsilon_5 \\ \underline{J}^{(2)} &= \underline{J}_S^{(2)} + \Upsilon_6\,\underline{L}^{(1)} + j_{10}\Upsilon_5^2 \\ \underline{L}^{(2)} &= \Upsilon_7\,\underline{L}^{(1)} + j_{20}\Upsilon_5^2 \quad , \end{aligned}\right\}(9.4)$$

where

$$\left.\begin{aligned} j_{10} &= -\tfrac{1}{4}\bar{\alpha}_1 u'_0 \\ j_{20} &= (\bar{\alpha}_2 - \tfrac{3}{4}\alpha_0^2\bar{\alpha}_1)y_0 + 2\alpha_0 j_{10} \qquad , \\ \text{with} \quad \bar{\alpha}_i &= \bar{N}\alpha_i \quad , (i = 1,2) \; . \end{aligned}\right\} \tag{9.5}$$

(b) The γ_i may now be rewritten in terms of i, i', u, u', y (in the notation of §20(b)). For the γ_i, (i= 1,2,3,4), we have as before

$$\left.\begin{aligned} \gamma_1 &= \tfrac{1}{2}(ii' + u'^2) \\ \gamma_2 &= \tfrac{1}{2}(u - u'^2) \\ \gamma_3 &= \tfrac{1}{4}(3ii' + 2u'^2 - 3u^2) \\ \gamma_4 &= -\tfrac{1}{4}(u'^2 - 4uu' + 3u^2) \quad , \end{aligned}\right\} \tag{9.6}$$

whilst for the new γ_i , (i=5,6,7), we have

$$\left.\begin{aligned} \gamma_5 &= y^2 \\ \gamma_6 &= k^2 yi + (\tfrac{3}{4} - k)\alpha_0\gamma_5 \\ \gamma_7 &= \alpha_0 yi + \gamma_1 \quad . \end{aligned}\right\} \tag{9.7}$$

It may be noticed that the γ_i have been so arranged that certain simple quantities recur several times, whereby computation is greatly simplified. Thus, for instance, the $\gamma_5,\gamma_6,\gamma_7$ are made up essentially merely of $\underline{y^2}$ and $\underline{yi}$. Notice that identities such as

$$\gamma_3 = \tfrac{3}{2}\gamma_1 + \gamma_4 - uu' \quad , \tag{9.81}$$

(uu' having been found in calculating γ_4), or

$$\gamma_1 - \frac{k}{(1-k)}\gamma_2 = \tfrac{1}{2}\alpha_0^2\gamma_5 \quad , \tag{9.82}$$

$$\text{and} \quad \gamma_7 = \gamma_3 + \gamma_4/(k-1) + \tfrac{3}{4}\alpha_0^2\gamma_5 \tag{9.83}$$

may be used as checks or otherwise.

Now $\gamma_1 = \tfrac{1}{2}(i_p i'_p + u'^2_p)\xi_1 + (i_p i'_q + u'_p u'_q)\eta_1 + \tfrac{1}{2}(i_q i'_q + u'^2_q)\zeta_1$ ...(9.9) and similarly for the other γ_i; whilst

$$\left.\begin{aligned} \underline{I} &= i_p\underline{Y}_1 + i_q\underline{V}_1 \\ \underline{Y} &= y_p\underline{Y}_1 + y_q\underline{V}_1 \quad . \end{aligned}\right\} \tag{9.91}$$

so that (4.2), (9.4-7) now provide the required polynomials $\underline{\Gamma}_{n,j}(Y_1, Z_1, V_1, W_1)$, $(n=1,2)$, (v. S §5(a)).

(c) If primary aberrations alone are considered great simplifications naturally arise, and moreover the aberrations may be written in a form closely resembling the Seidel sums (v. S §16). For further details an earlier paper by the author may be consulted. (Buchdahl, 1948)

§10. Special surfaces.

(a) The plane surface does not require any special treatment. All that is necessary is to put α_0 equal to zero in the formulae of the preceding section. However, in view of the simplicity of the consequent results we write the final equations in terms of only three γ's, viz.

$$\left.\begin{aligned} \gamma_1 &= \tfrac{1}{2}(1-k^2)\zeta \\ \gamma_2 &= \xi \\ \gamma_3 &= -k^2\eta \quad , \end{aligned}\right\} \quad (10.1)$$

so that
$$\left.\begin{aligned} \underline{J}_s^{(1)} &= N y_0 \gamma_1 \\ \underline{J}_s^{(2)} &= -\tfrac{3}{2}\gamma_1 \underline{J}_s^{(1)} \quad . \end{aligned}\right\} \quad (10.21)$$

Hence
$$\left.\begin{aligned} \underline{J}^{(1)} &= \underline{J}_s^{(1)} \\ \underline{L}^{(1)} &= \bar{\alpha}_1 y_0 \gamma_2 \\ \underline{J}^{(2)} &= \underline{J}_s^{(2)} + \gamma_3 \underline{L}^{(1)} + j_{10}\gamma_2^2 \\ \underline{L}^{(2)} &= \bar{\alpha}_2 y_0 \gamma_2^2 + \alpha_1 \underline{J}^{(1)} \gamma_2 \quad . \end{aligned}\right\} \quad (10.22)$$

(b) By definition the figured sphere implies $\alpha_1 = 0$, $\alpha_0 \neq 0$. In that case (9.4) reduces to

$$\left.\begin{aligned} \underline{J}^{(1)} &= \underline{J}_s^{(1)} \quad , & \underline{L}^{(1)} &= 0 \\ \underline{J}^{(2)} &= \underline{J}_s^{(2)} \quad , & \underline{L}^{(2)} &= \bar{\alpha}_2 y_0 \gamma_5^2 \quad . \end{aligned}\right\} \quad (10.3)$$

It will be seen that only a single new term need be computed, as compared with the sphere. γ_6 and γ_7 are here redundant. The simplicity of (10.3) shows that the effect of figuring the surfaces (in our sense) of a system originally containing only spherical surfaces is very easily dealt with, to this order. In practice this problem is likely frequently to occur.

(c) (i) The general expressions derived above may be partly checked (though fairly reliably so) by considering an optical system consisting of a single refracting surface which possesses a pair of conjugate points on its axis of symmetry; in the sense that all the members of a congruence of rays which pass through the first point also pass through the second after refraction, either actually or virtually. For all such rays the expression (4.2) must then vanish identically, i.e. we must have

$$\Delta\underline{\Lambda}^{(n)} \equiv 0 \quad , (n = 1,2,\ldots) \ . \tag{10.4}$$

The curve of intersection of such a surface with the $\underline{x}$, $\underline{y}$ plane is, in general, a Cartesian Oval, so that its equation is

$$\Delta N\sqrt{(x-a)^2 + \chi} = \Delta Na \quad , \tag{10.5}$$

where $\underline{a}$, $\underline{a}'$ are the $\underline{x}$-coordinates of the axial conjugate points.

Taking the surface to be described in terms of α_0, $\underline{a}$, $\underline{k}$ we have

$$1/a' = (1-k)\alpha_0 + k/a \quad . \tag{10.51}$$

If we substitute the series (6.1) in (10.4) we then find

$$\left.\begin{aligned} \alpha_1 &= -\tfrac{1}{2}k(\alpha_0 - \tfrac{1}{a})^2(k\alpha_0 - \tfrac{1+k}{a}) \\ \alpha_2 &= -\tfrac{3}{4}(\alpha_0 - \tfrac{1}{a})(k^2\alpha_0 - \tfrac{1+k+k^2}{a})\alpha_1 \quad . \end{aligned}\right\} \tag{10.6}$$

The vanishing of the second factor of α_1 represents the case in which the ovoid degenerates into a sphere, the two conjugate points then being the usual aplanatic points of the sphere. It follows incidentally the the ovoid can never be a figured sphere. If $a \to \infty$ the surface degenerates into a prolate ellipsoid (spheroid) of eccentricity k.

(ii) Since the object point lies on the axis we need not distinguish between y_0, u_0 and y, u. We now have $y = au$, and $\alpha_0 y = 1+u$. For α_1 and α_2 we may therefore write

$$\left.\begin{aligned} \alpha_1 &= -k1^2(1-u')/2a^3u^3 \\ \alpha_2 &= -3\alpha_1 1(1-u'-ku')/4a^2u^2 \,. \end{aligned}\right\} \quad (10.7)$$

In virtue of these relations it may be verified that $\Delta\Lambda^{(1)}$ and $\Delta\Lambda^{(2)}$ vanish identically.

§11. On the design of systems with aspherical surfaces.

(a) The problem considered in S consisted solely on determining the aberrations of a given optical system in the form of power series in certain variables. Moreover it is clear that the coefficients of these series are exceedingly complex functions of the parameters defining the system, i.e. of N_j, r_j, d'_j. It follows that if in the course of design finite changes be made in any of these parameters the whole computation must be carried out again with the new values, or, at any rate for all surfaces following the i-th, if no changes have been made on the 1st, 2nd,..., i-th surfaces.

When we come to aspherical surfaces however, the position is radically different. For if we consider, for the sake of argument, only primary and secondary aberrations there are now at our disposal $2k$ new degrees of freedom for a system with k surfaces, viz. the asphericities θ_{1j}, θ_{2j}, $(j=1,2,\ldots,k)$; and any of these asphericities can be altered at will without any consequent change in the paraxial coefficients of the system.

Moreover if we use the equations set out in the previous sections we may compute the aberrations of the system using chosen values of the N_j, r_j, d'_j ,$(j=1,2,\ldots,k)$. For in (9.4) the coefficients of α_1 and α_2 are functions of the latter alone, and are therefore readily calculated. Now, α_1 and α_2 enter linearly into the expressions (9.4), and therefore linearly into $\underline{\Delta\Lambda}$ when the increments are neglected. But the increments affect only the secondary aberrations, the primary terms of the former alone contributing. From which follows the important result that <u>the primary terms of ε'_k are linear functions of the θ_{1j} and independent of the θ_{2j} , whilst the secondary terms of ε'_k are linear functions of the θ_{2j} and quadratic in the θ_{1j}.</u> $(j=1,2,\ldots,k)$.

(b) It should be realised that even when the total number of refracting surfaces is not restricted it is not in general possible to reduce to zero all the primary and secondary aberration coefficients (for a given position of the object) by a suitable choice of of the θ_{1j}, θ_{2j}. Thus, in particular, the primary coefficients can be reduced to zero only if the "paraxial constitution" of the system has been so chosen that the Petzval sum vanishes, and analogous conditions will hold for the higher order coefficients. But naturally it will usually be possible to obtain a much better compromise as regards the balance of the final aberrations.

We may proceed, then, by leaving the θ_{1j} and θ_{2j} as unknown quantities in the computations, subsequently to be determined in accordance with the demands imposed ab initio upon the magnitude of the final aberrations, (i.e. by practical considerations). But if something is already known about the general behaviour of the system it may be simpler to proceed as follows. The primary aberrations are computed first, with unknown θ_{1j}. The latter are then given definite values in some way. The secondary terms are then computed with unknown θ_{2j}; and then the latter are determined. This procedure has the advantage that at every stage the asphericities enter only linearly into the expressions to be considered in determining them.

(c) In the equations of §9(a) the division into two sets of terms suggested itself quite naturally, viz. the set of terms which vanish when the surface is spherical, and the set of those which do not. This has special advantages from a practical viewpoint insofar as systems with initially only spherical surfaces will frequently be taken as a basis for subsequent design. And something will in general already be known about the aberrations of this system,-either from ray traces or otherwise,- which may be of considerable help in deciding the values of the θ_{1j} when the course described above under (b) is adopted. Moreover if, (in the case of such a 'base system') we decide upon only figuring the surfaces (i.e. leaving $\alpha_{1j}=0$), then in view of the results of §10(b) the only new term which appears in $\Delta\underline{\Lambda}$ is $\bar{\alpha}_2\xi^2 y_0\underline{Y}$. (Special provision will have to be made for plane surfaces, unless we agree to leave α_1 zero, i.e. to figure it by altering α_2 only). It follows that if for any ray $(\underline{Y}_1,\underline{V}_1)$ the original aberration was $(\underline{\varepsilon}_k')_s+\underline{O}(7)$, then after figuring it will be given explicitly by

$$\underline{\varepsilon}'_k = (\underline{\varepsilon}'_k)_s + \sum_{j=1}^{k} \frac{y_{oj}}{N'_k u'_{ok}} \bar{a}_{2j}(y_p^2\xi_1 + 2y_py_q\eta_1 + y_q^2\zeta_1)_j(y_p\underline{Y}_1 + y_q\underline{V}_1)_j + O(7). \quad \ldots(11.1)$$

A few lines of calculation therefore give all the information required concerning the effect of simultaneously figuring some or all of the surfaces, to the order considered. The $\bar{a}_{2j}$ may of course again be left as unknowns, to be determined subsequently as the solution of a set of linear simultaneous algebraic equations.

(d) It will be noticed that paraxial coefficients alone enter into the equation (11.1) for the difference $\delta\underline{\varepsilon}' \equiv \underline{\varepsilon}'_k - (\underline{\varepsilon}'_k)_s$; so that in practice we might deal with the problem of figuring the surfaces of a base system by means of the following hybrid procedure. We first find $(\underline{\varepsilon}'_k)_s$ for a number of selected rays by means of ray traces. We then calculate $\delta\underline{\varepsilon}'$ from (11.1) for these rays. Then provided tertiary and higher aberrations are reasonably small compared with the primary and secondary the resultant value of $\underline{\varepsilon}'_k$ will be a good approximation to the true value. This method has the virtue of great simplicity, especially as ray traces of the base system will generally already be available.

(e) The 'natural' separation of (9.4), to which we referred under (c) above, into two sets of terms especially adapted to the case of 'spherical' base systems arose essentially because of the particular choice of constants (viz. a_n) used in the specification of the shape of the surface. If practical requirements ever justified such a course we could, however, express the shape of the surface alternatively in terms of some set of coefficients, σ_n say, in such a way that σ_m, $(m > 0)$, would imply a desired 'base surface'. For example, if we chose σ_m to be θ_{m+1} the base surface would be a paraboloid of revolution, and so on.

(f) In this section we have so far taken only primary and secondary aberrations into account explicitly. But the extension to higher orders is immediately apparent. Thus, for instance, the values of the coefficients θ_{nj} affect only the aberrations of order $\geqslant n-1$; and in particular the aberrations of order n depend linearly on the $\theta_{n+1,j}$, quadratically on the $\theta_{n,j}$; and so on. In the same way the considerations of sections (c) and (d) above may be extended to higher orders. If the aberrations of order $n+1$ are negligible, but those of order n are not, then use may be made of an equation analogous to (11.1) to describe the change $\delta \underline{\varepsilon}'$ consequent upon introducing asphericities α_m into a spherical base system, the α_m, $(0 < m < n)$, retaining their value zero.

§12. Example of the application of the method.

(a) No suitable design of a system with aspherical surfaces was available for the purpose of applying the method here developed to the analysis of aberrations of such a system. A more far reaching problem was therefore attacked, *viz.* the further improvement of the 'corrected' Cooke Triplet of S§25, after the manner described in the previous section. Since the author, working on his own, wished as far as possible to avoid further algebraic computation it was decided amongst the θ_{2j} to change only $\theta_{2,6}$ since the latter does not of course enter into the primary terms of the increments $\underline{\delta}_{yj}$, $\underline{\delta}_{vj}$, $(j=1,\ldots,6)$. Arbitrary changes of the θ_{3j} were however to be allowed. In this way the available computations of the primary and secondary aberration coefficients of the original system could be made use of in their entirety (v.§11(c)). Also, it was decided to change θ_{3j} of as few surfaces as possible and of surfaces occurring as late as possible in the system in order to minimise the labour of obtaining trigonometrical comparison traces; for which reason also attention was mainly concentrated on tangential rays. It must be understood that these arbitrary restrictions make this example somewhat trivial in the sense that it illustrates only to a small degree the power of the present methods. We emphasise especially that in actual practice such restrictions would not in general exist nor would their absence bring about any additional difficulties in, or appreciable lengthening, of the computations. In fact, computing schemes will be very similar to that of S§25(b), and the remarks made in that section as regards number of entries per surface etc. remain generally valid, and indeed gain added strength if the labour involved in tracing skew rays through aspherical surfaces be considered.

(b) The adjustment of $\theta_{2,6}$ merely requires the calculation at the sixth surface of the coefficients of $\underline{L}^{\omega}\underline{Y}$. Remembering that $N'_k u'_{ok} = 1$ we write, omitting the subscript 6,

$$\underline{\varepsilon}' = A'\underline{Y}_1\xi_1 + \bar{A}'\underline{V}_1\xi_1 + B'\underline{Y}_1\eta_1 + \bar{B}'\underline{V}_1\eta_1 + C'\underline{Y}_1\zeta_1 + \bar{C}'\underline{V}_1\zeta_1 + \underline{Q}(5). \quad \ldots(12.1)$$

where

$$\left.\begin{aligned} A' &= +1.35920 - 0.300986\alpha_1 \\ \bar{A}' &= -0.168567 + 0.083432\alpha_1 \\ B' &= -0.337133 + 0.166864\alpha_1 \\ \bar{B}' &= +0.009739 - 0.046254\alpha_1 \\ C' &= +0.174952 - 0.023127\alpha_1 \\ \bar{C}' &= -0.035468 + 0.006411\alpha_1 \quad . \end{aligned}\right\} \quad (12.11)$$

(Notice that the object is again taken to be at infinity.)

Knowing that the spherical aberration is a dominating factor amongst the aberrations as a whole we now aim at leaving a small positive residual primary spherical aberration. Accordingly we choose

$$\alpha_1 = +2.8 \quad , \qquad (12.2)$$

or, what comes to the same thing,

$$\theta_2 = +0.310053 \quad , \qquad (12.21)$$

compared with its original value of $\theta_2 = -0.389947$. This tends to reduce considerably also the other large terms, especially as regards the tangential aberration coefficients. The surface is therefore approximated in the extended paraxial region by the hyperboloid of revolution

$$\frac{(X - 0.86074)^2}{0.74087} - \frac{Y^2 + Z^2}{0.58908} = 1 \quad , \qquad (12.22)$$

the eccentricity of which is 1.340.

Since now $\alpha_{1,6} \neq 0$, certain extra terms appear in the $\underline{J}^{(2)}$, $\underline{L}^{(2)}$ at this surface, which are absent at the others. The effect of these on the final secondary aberrations is very small, as may be seen by comparing the first column of the table of section (c) below with the secondary coefficients of the series $\underline{S}(25.2)$.

(c) The result of altering the $\alpha_{2,j}$ may now be computed on the basis of (11.1); it is given in the following table.

	0	1	2	3	4	5	6
S'_1	-95.895	$+3.6972$	-2.3297	$+1.4749$	-1.3210	$+1.3062$	-1.2621
$\bar{S}'_1$	-13.234	0	$+.06270$	$-.07060$	$+.07414$	$-.32605$	$+.34986$
S'_2	-50.839	0	$+.25082$	$-.28240$	$+.29656$	-1.3042	$+1.3994$
$\bar{S}'_2$	$+5.0796$	0	$-.0_2675$	$+.01352$	$-.01664$	$+.32556$	$-.38792$
S'_3	$-.90274$	0	$-.0_2338$	$+.0_2676$	$-.0_2832$	$+.16278$	$-.19396$
$\bar{S}'_3$	$+1.6731$	0	$+.0_4908$	$-.0_3323$	$+.0_3467$	$-.04063$	$+.05376$
S'_4	$+4.4778$	0	$-.0_2675$	$+.01352$	$-.01664$	$+.32556$	$-.38792$
$\bar{S}'_4$	$+1.4363$	0	$+.0_3182$	$-.0_3647$	$+.0_3934$	$-.81266$	$+.10753$
S'_5	$+3.1727$	0	$+.0_3182$	$-.0_3647$	$+.0_3934$	$-.31266$	$+.10753$
$\bar{S}'_5$	$-.79845$	0	$-.0_5489$	$+.0_4310$	$-.0_4524$	$+.02029$	$-.02981$
S'_6	$-.69777$	0	$-.0_5122$	$+.0_5774$	$-.0_4131$	$+.0_2507$	$-.0_2745$
$\bar{S}'_6$	$+.03964$	0	$+.0_7329$	$-.0_6371$	$+.0_6736$	$-.0_2127$	$+.0_2207$

... (12.3)

Thus, let $\sigma_j = (\theta_3 - 2\theta_1^2\theta_2)_j = (\frac{1}{6}\alpha_2 + \frac{1}{8}\alpha_0^2\alpha_1)_j$; (12.31)

and let μ_{ij} be the entry in the $\underline{j}$-th column and the row marked S'_i . Then, with $\sigma_0 \equiv 1$,

$$S'_i = \sum_{j=0}^{6} \sigma_j \mu_{ij} \quad , \tag{12.4}$$

and similarly for the $\bar{S}'_i$, where the S'_i, $\bar{S}'_i$ are the secondary aberration coefficients, (i.e. as in $\underline{S}(3.93)$, with the index $\underline{k}$ omitted). From the table it will be seen that the effect of figuring the 3rd surface is much the same as that of figuring the 4th, so that we need only consider the latter. The 5th and 6th surfaces are equivalent in the same sense, so that we only consider the sixth. The 1st and 2nd again are equivalent, but in accordance with the remarks made under (a) above it was decided to alter neither of these. Effectively, therefore, only the 4th and 6th surfaces remain to be considered. Inspection of the table shows the values

$$\sigma_4 = -70 \quad , \quad \sigma_6 = +15 \tag{12.5}$$

to be a fairly reasonable compromise, and these were the values adopted. The corresponding values of α_2 and θ_3 are

$$\left.\begin{aligned} \alpha_{2,4} &= -420 \quad , & \alpha_{2,6} &= +85.5174 \\ \theta_{3,4} &= +149.458 \quad , & \theta_{3,6} &= +15.3311 \quad . \end{aligned}\right\} \tag{12.6}$$

Accordingly we take as the equations of the 4th and 6th surfaces

$$\left.\begin{aligned} &\text{4th: } x = 2.55897\chi + 16.7569\chi^2 + 149.459\chi^3 \\ &\text{6th: } x = -0.730581\chi + 0.310053\chi^2 + 15.3311\chi^3 \quad . \end{aligned}\right\} \tag{12.7}$$

For the complete primary and secondary aberrations we then have

$$\begin{aligned} \underset{\sim}{\varepsilon}'_6 = &+0.56144\underline{Y}_1\xi_1 + 0.065042\underline{V}_1\xi_1 + 0.13008\underline{Y}_1\eta_1 - 0.11977\underline{V}_1\eta_1 \\ &+0.11019\underline{Y}_1\zeta_1 - 0.01752\underline{V}_1\zeta_1 \\ &-22.357\underline{Y}_1\xi_1^2 + 13.176\underline{V}_1\xi_1^2 - 50.606\underline{Y}_1\xi_1\eta_1 + 0.42579\underline{V}_1\xi_1\eta_1 \\ &-3.2296\underline{Y}_1\xi_1\zeta_1 + 2.4469\underline{V}_1\xi_1\zeta_1 - 0.17600\underline{Y}_1\eta_1^2 + 2.9838\underline{V}_1\eta_1^2 \\ &+4.7203\underline{Y}_1\eta_1\zeta_1 - 1.3345\underline{V}_1\eta_1\zeta_1 - 0.80855\underline{Y}_1\zeta_1^2 + 0.070575\underline{V}_1\zeta_1^2 \quad . \end{aligned}$$

$$\cdots (12.8)$$

§13. Comparison of results with trigonometrical traces.

(a) Because of the very great labour involved in tracing an appreciable number of skew rays through aspherical surfaces, trigonometrical comparison traces were obtained only for tangential pencils of rays, for which (12.8) reduces to

$$\begin{aligned}\varepsilon'_y = &+0.51644Y_1^3 + 0.19512Y_1^2V_1 - 0.00957Y_1V_1^2 - 0.017518V_1^3 \\ &- 22.357Y_1^5 - 63.782Y_1^4V_1 - 2.9798Y_1^3V_1^2 + 10.151Y_1^2V_1^3 \\ &- 2.1430Y_1V_1^4 + 0.070575V_1^5 \quad . \end{aligned} \tag{13.0}$$

The following point may be noticed: We have taken (12.7) to be the exact representation of the 4th and 6th refracting surfaces, i.e. we adjusted all the higher order ashpericities of these two surfaces by setting $\theta_{n,4} = \theta_{n,6} \equiv 0$, $(n \geqslant 4)$. And all trigonometrical traces were computed on the basis of (12.7).

(b) From (12.8) we obtain the aberration ε' for any given ray. In practical optics, however, it is usual to speak of distortion coma, etc., in the actual image although these terms have a simple significance strictly speaking only in the context of primary aberrations. It may therefore be useful sometimes to adopt definitions of certain types of aberrations which correspond closely to practical usage. For example, we may define distortion corresponding to a certain pencil of rays to be the value of ε' for the principal ray of that pencil. Now complications will arise from the fact that the image of the diaphragm formed by that part of the optical system which precedes it is not in general perfect; that is, there is no definite entrance pupil. But, clearly, for practical purposes these and similar difficulties may be overcome simply by taking the principal ray of a pencil to be the ray which passes through the centre of the <u>paraxial entrance pupil</u>, (which we shall take to lie at a distance <u>p</u> behind the first surface). The principal ray so defined will in general not pass precisely through the centre of the diaphragm, but this is obviously of no practical consequence. Let this ray have the coordinates $(Y_1,V_1,Z_1,W_1) \equiv (\bar{Y},\bar{V},0,0)$, the object point having been taken to have the coordinates $(l_{01}, -h, 0)$, without loss of generality. Then

$$\left.\begin{aligned} \bar{Y} &= p\bar{V} \quad , \\ \text{where} \quad \bar{V} &= h/(l_{01} - p) \quad . \end{aligned}\right\} (13.1)$$

Then the coordinates of any other member of the pencil of rays are given in terms of those of the principal ray by means of polar coordinates σ, ψ , <u>viz</u>.

$$\left.\begin{aligned} Y_1 &= \bar{Y} + \sigma\cos\psi \quad , \\ Z_1 &= \sigma\sin\psi \\ V_1 &= (Y_1 + h)/l_{01} \\ W_1 &= Z_1/l_{01} \quad . \end{aligned}\right\} (13.11)$$

If the object is at infinity, as in our case, these reduce to

$$\left.\begin{aligned} &\bar{Y} = p\bar{V} \quad , \quad \bar{V} = \text{const.} \\ \text{and} \quad &Y_1 = \bar{Y} + \sigma\cos\psi \quad , Z_1 = \sigma\sin\psi \; , V_1 = \bar{V} \; , W_1 = 0. \end{aligned}\right\} (13.12)$$

We have implied that the aperture of the diaphragm is circular. [This restriction, of course, is quite unessential]. The radius of this is to be thought of as determining the corresponding value of σ according to paraxial laws, [<u>cf.</u> the author's earlier paper (Buchdahl, 1946), §§ 3(b) and 10-13. The quite unnecessary complications arising from not adopting the simple definitions of this section are there very apparent].

For the distortion we now simply have (in our case)

$$\text{distortion} = D_1\overline{V}^3 + D_2\overline{V}^5 + O(7) , \qquad (13.2)$$

with $$D_1 = \sum_{i=1}^{3} a_i p^i \quad , \quad D_2 = \sum_{i=1}^{6} b_i p^i \quad , \qquad (13.21)$$

where a_i and b_i are respectively the coefficients in (13.0) of $Y_1^i V_1^{3-i}$ and $Y_1^i V_1^{5-i}$.

Since it is doubtful whether aberrations defined in this manner are of much use in more advanced applied optics we confine our further considerations to tangential coma, since the latter often gives fairly reliable information concerning the asymmetrical distribution of light in the image. We define it, for any object position, by the equation

$$\text{tang. coma.} = \Gamma = \tfrac{1}{2}\left[\varepsilon'(\overline{Y}+\sigma) + \varepsilon'(\overline{Y}-\sigma)\right] - \varepsilon'(\overline{Y}) , \qquad (13.3)$$

for the case of a definite pencil of semi-aperture $(\alpha)_1\sigma$ at incidence. (13.3) is an adaptation of the definition of Conrady, (Conrady, 1929). When the object is at infinity it reduces to

$$\Gamma = \frac{1}{2!}\frac{d^2D_1}{dp^2}\overline{V}\sigma^2 + \frac{1}{2!}\frac{d^2D_2}{dp^2}\overline{V}^3\sigma^2 + \frac{1}{4!}\frac{d^4D_2}{dp^4}\overline{V}\sigma^4 + O(7) . \qquad (13.31)$$

Hence if (in this case) attention is paid chiefly to the removal of spherical and tangential coma the asphericities must be adjusted so that the five conditions

$$a_3 = a_2 = b_5 = b_4 = 3b_3p + b_2 = 0 \qquad (13.4)$$

are satisfied as nearly as possible.

If we define sagittal coma in an analogous fashion, the aberration characterised by the two coma coefficients might be called 'total coma', for the total coma is then the resultant of various types of coma, viz. circular coma of various orders, elliptical coma of various orders, etc.

(c) The diagrams 1. - 6. graphically represent a comparison between aberrations as predicted on the basis of (13.1), and the aberrations as obtained from a considerable number of strict trigonometrical traces. All measures are given in % focal length.

(i) Figures 1. - 4. In these the aberration ε' $(=\varepsilon'_y)$ is plotted against Y_1 for the case of four groups of rays making angles of 0^o (spherical aberration), 6^o, 12^o, 18^o respectively with the axis of the system in the object space. The meaning of the curves is as follows:

.... primary aberrations,

- - - - - - - primary + secondary aberrations,

———— full aberrations given by trigonometrical traces.

These curves may be compared with figures 1.,3.,4., and 5. of the paper quoted above (Buchdahl, 1946), if due attention is paid to the fact that the coordinate y used there is equivalent to the present $Y_1/\sqrt{1+Y_1^2}$, whilst y' is equivalent to $0.9673\varepsilon'$. It will be seen that the predicted aberrations agree well with those found by strict trigonometrical ray tracing.

Moreover, the aberrations are on the whole much improved just in the way aimed at in choosing the particular asphericities which were introduced. It is interesting to note how much is achieved by means of the very restricted changes made. In

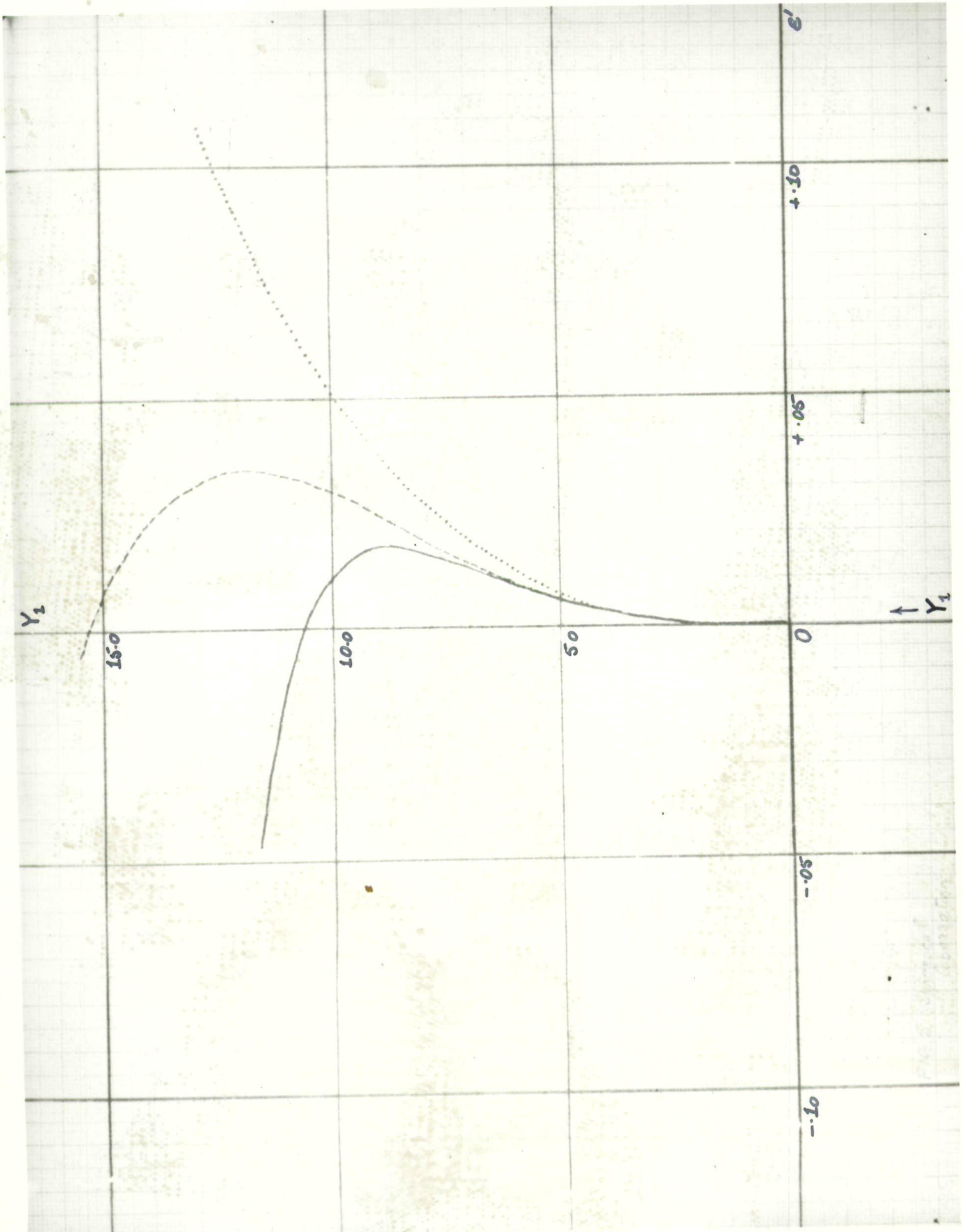


Figure 1. 0^{o} pencil (spherical aberration).

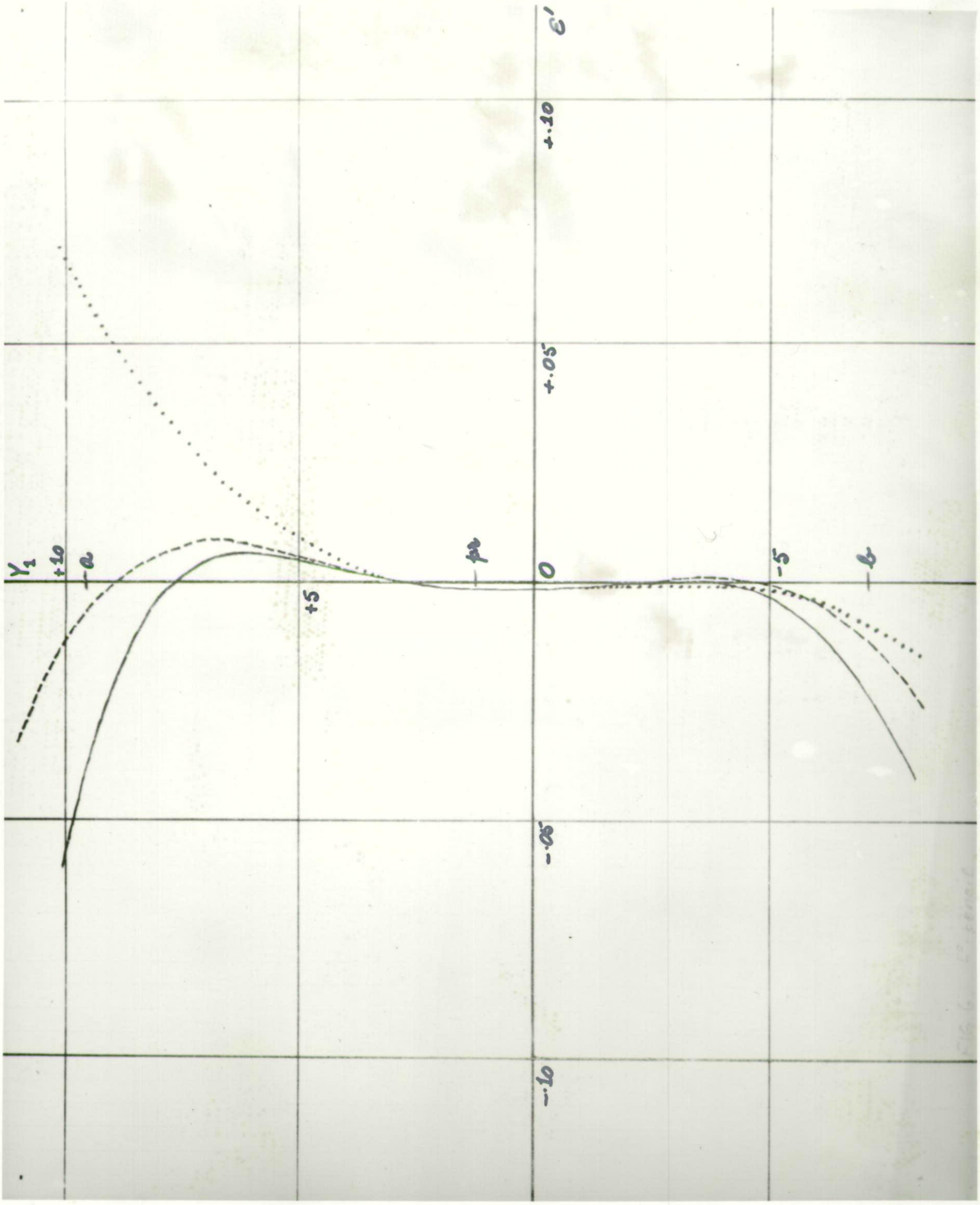


Figure 2. 6° pencil.

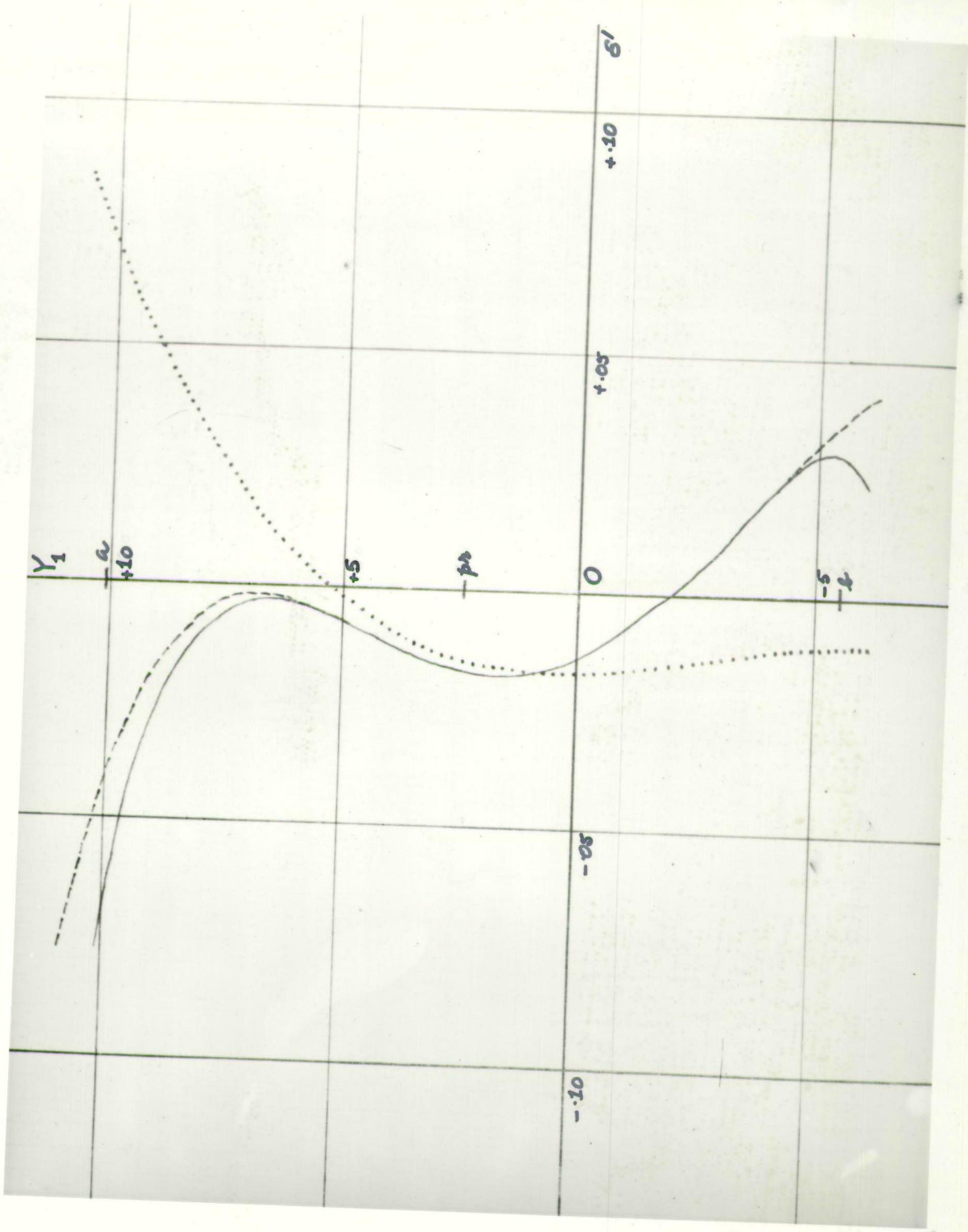


Figure 3. 12° pencil.

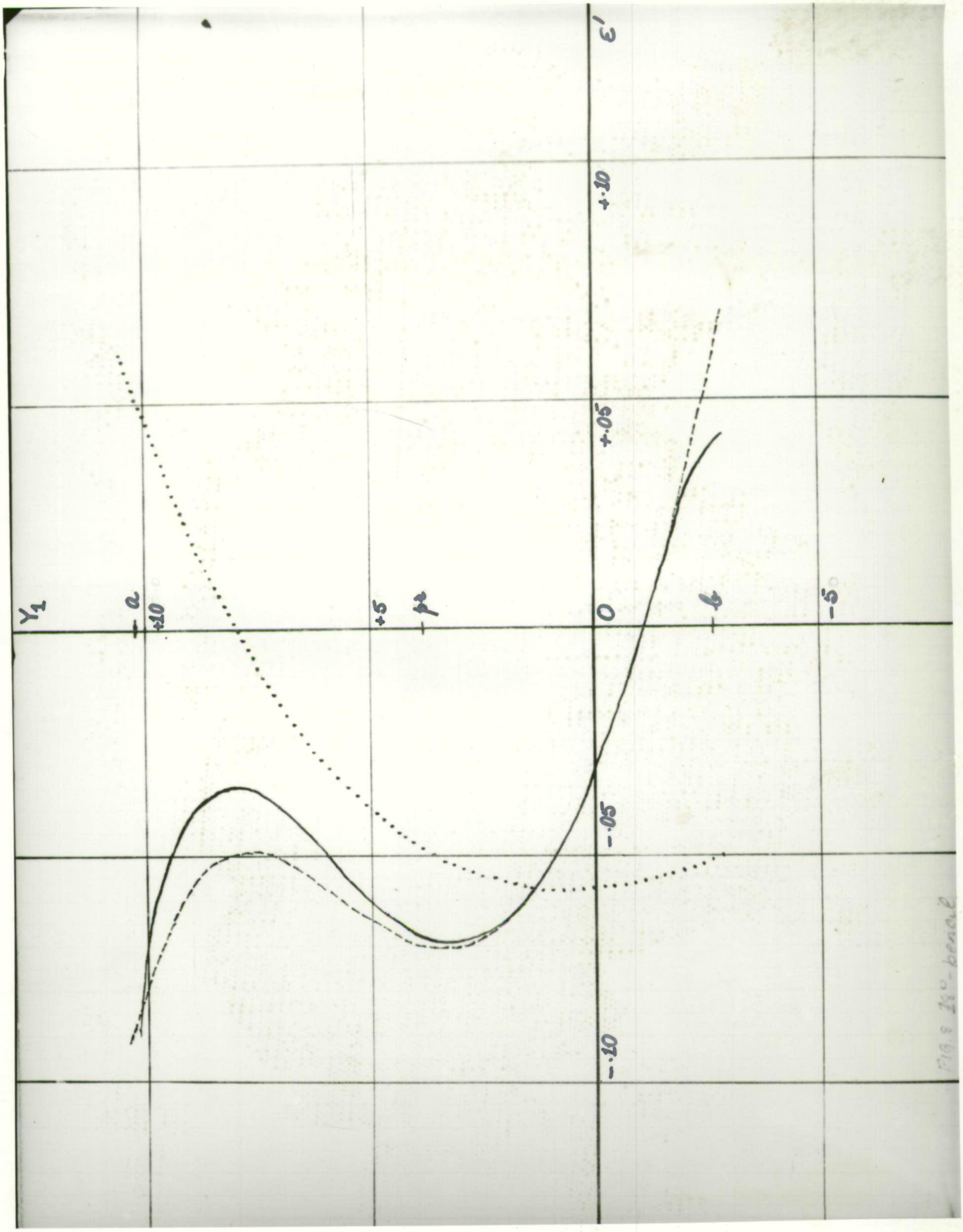


Figure 4. 18° pencil.

this context figure 5. may also be considered.

(ii) Figure 5. graphically represents the aberrations for the 6^{O} pencils of the original system (curve s.), and of the altered system (curve a.), both drawn on the same scale. They are as obtained from a number of strict trigonometrical traces.

The curve s being available we could also determine the aberrations of the altered system according to the procedure described in §11(d). The corresponding curve lies very close to the actual curve a between $Y_1 = +10$ and $Y_1 = -7$ (which comprises the total range required), and coincides with it, in fact, between $Y_1 = +6$ and $Y_1 = -4$. (This means that for values of the stop-number down to about f/9 this very simple procedure gives rise to accurate predictions of aberrations, in the case of the present system).

The points marked a,b and pr respectively on the various diagrams refer to the extreme rays, and to the principal rays of the pencils with the diaphragm in the position which it occupied in the original system. If two rays have Y_1 coordinates such that $Y_{1a} - \bar{Y} = \bar{Y} - Y_{1b} = \sigma$ then we can always consider them to be the extreme rays of a pencil (i.e. the rays just grazing the rim of the aperture of the diaphragm). In the present case the effective stop-number is then almost exactly

$$\text{stop-number} = f/2\sigma \quad . \tag{13.5}$$

Speaking generally the position of the diaphragm, which is not involved in the computations for ε', may be determined directly

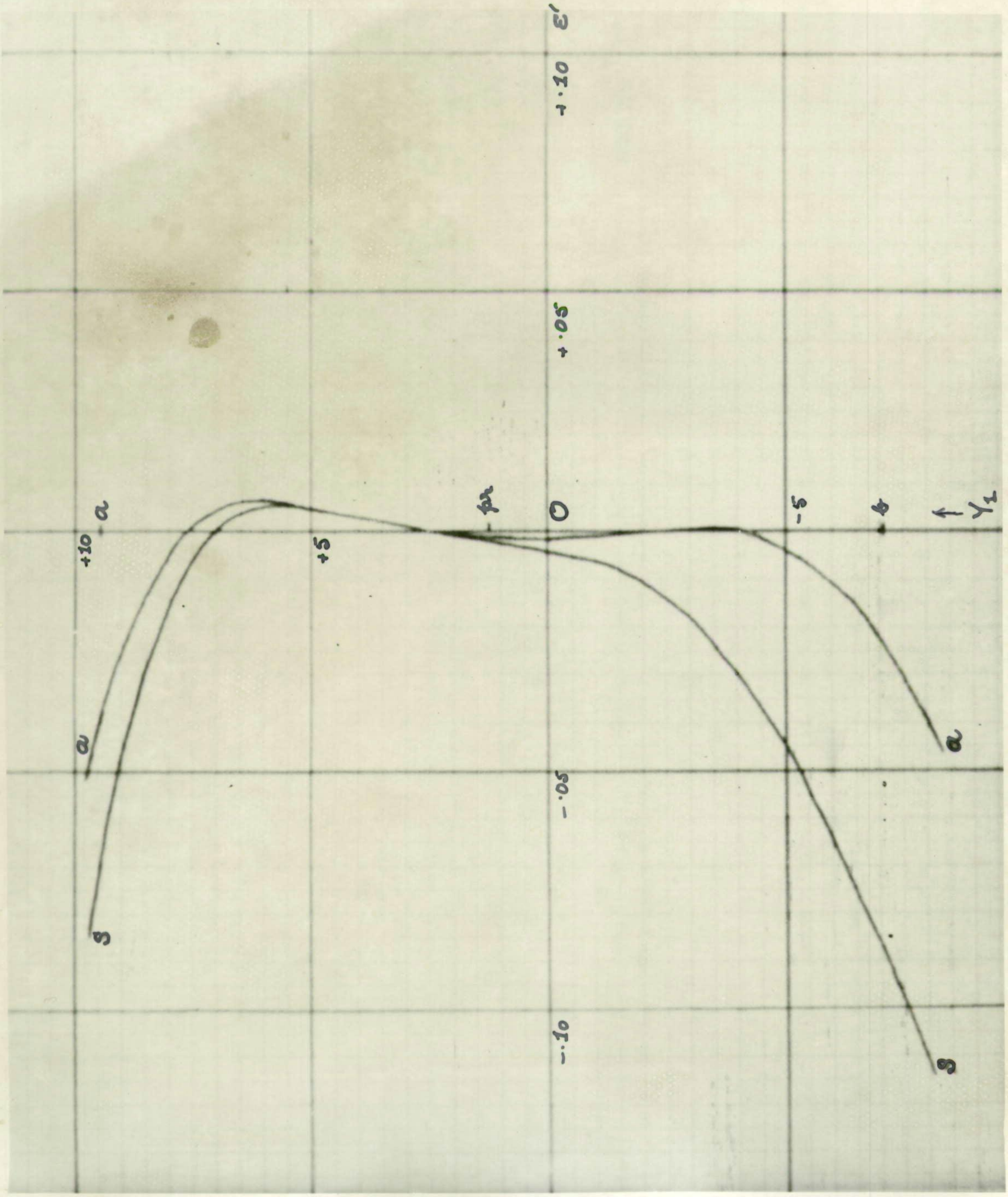


Figure 5. 6° pencils for 's' and 'a' systems.

from the latter. The criteria to be applied will vary from case to case. On this occasion the old diaphragm position may be retained. For the 6° diagram shows that if we choose Y_{1a}, Y_{1b} such that $|\varepsilon'| <$ some chosen value, say 0.02, then Y_{pr} falls nearly into the position indicated. However, for an adequate determination of the most favourable position a knowledge of the aberrations to this order will rarely be sufficient. Some qualitative knowledge, at least, of the tertiary aberrations will be required, and in practice this will often be available.

(iii) Figures 6(a) and 6(b). These represent the coma of the 6° and 12° pencils as a function of the aperture of the pencils, the latter being specified conveniently by σ (v. §13(b)). Still giving linear measures in terms of % focal length, we have

$$p = 11.3227\ , \tag{13.6}$$

whence (13.31) gives

$$\Gamma = +\,37.05\bar{V}\sigma^2 + (320.8\bar{V}^3\sigma^2 - 7644\bar{V}\,\sigma^4) + \underline{O}(7)\ , \tag{13.7}$$

which may be compared with the predicted coma of the unchanged system: $\Gamma = -4.712\bar{V}\sigma^2 + (307.3\bar{V}^3\sigma^2 - 14070\bar{V}\sigma^4) + \underline{O}(7)$. (13.71)

It will be seen that the primary coma is greatly increased. This increase is however more than compensated by a large decrease in secondary (circular) coma. The curves are as given under (i) above. The additional unbroken curves marked s represent the coma of the original system, given by strict ray traces, (cf. figures 8. and 9. of the paper quoted in (i).)

(iv) For the distortion we find

$$\text{dist.} = -1.535\bar{V}^3 - 4.497\bar{V}^5 + \underline{O}(7) \tag{13.8}$$

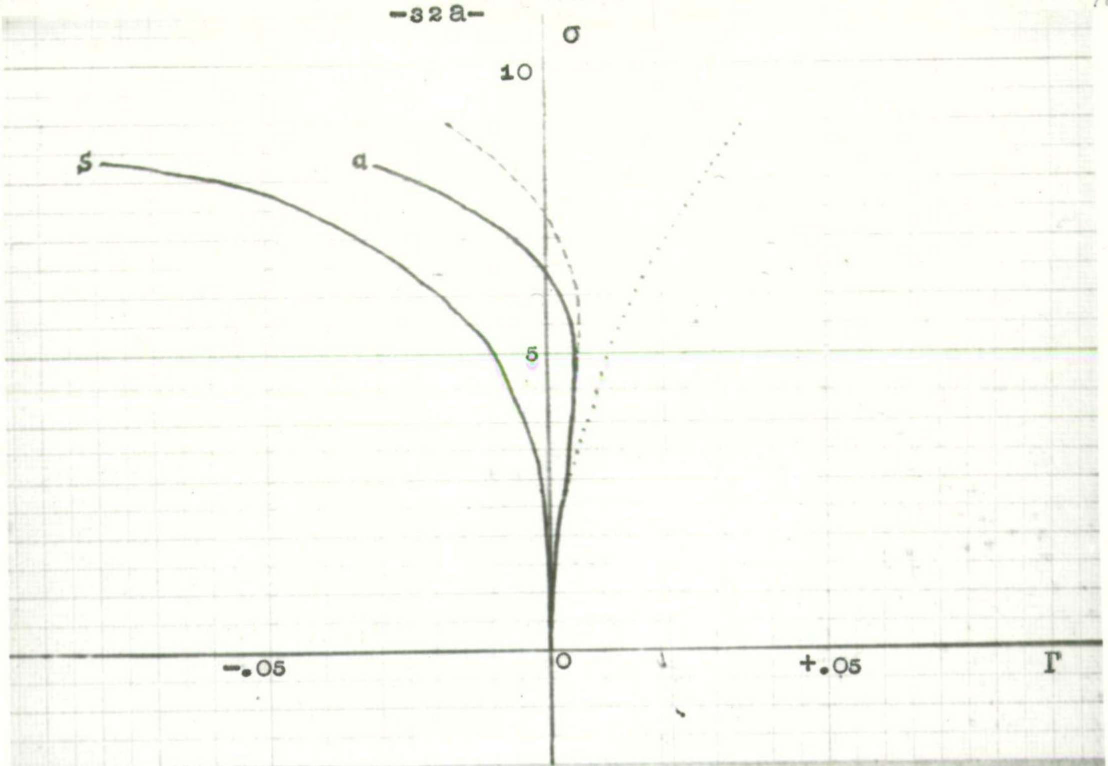


Figure 6(a).

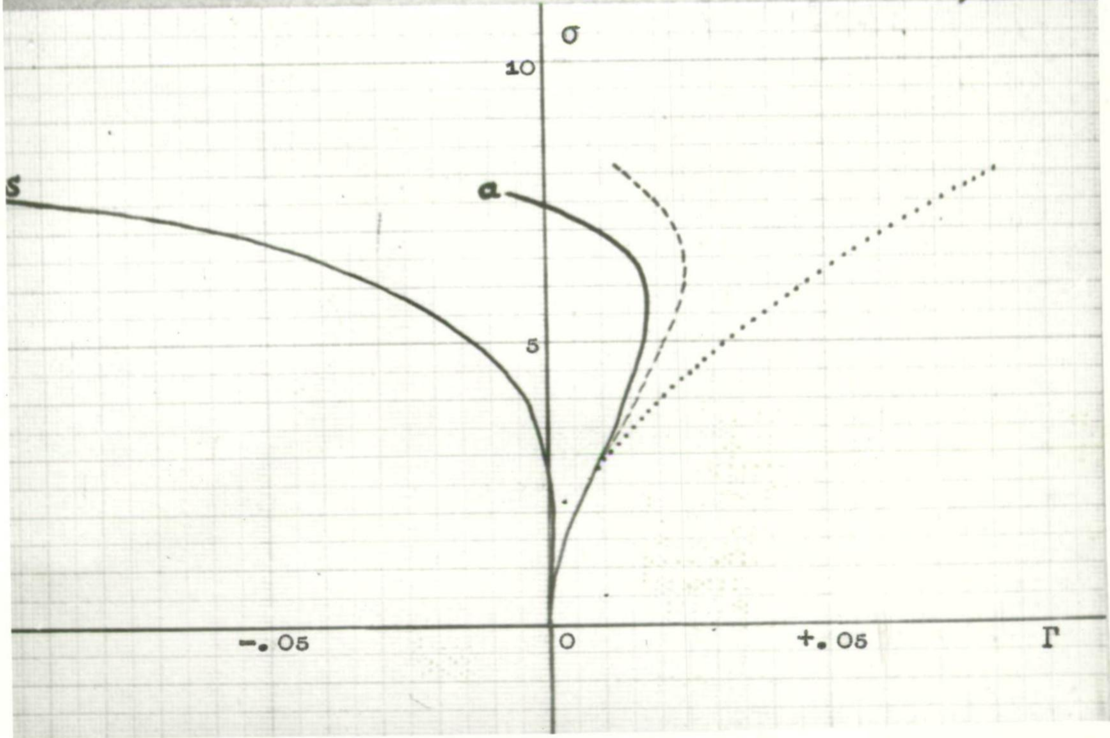


Figure 6(b).

Tangential coma as function of aperture, 6° and 12°.

compared with that of the original system

$$\text{dist.} = -1.907\bar{V}^3 - 6.1173\,\bar{V}^5 + \underline{O}(7) \;. \qquad (13.81)$$

The curves based on (13.8) have been omitted since the broken curve (i.e. primary + secondary aberrations) and the full curve (full aberrations) actually coincide on a diagram drawn on a scale similar to that of the other diagrams.)

The various curves clearly demonstrate that the asphericities may be adjusted, and the aberrations thereby suitably modified according to the predictions of the 'algebraic' aberration coefficients.

§14. Conclusion.

There is little point in repeating the remarks of S§27. But we do repeat that the amount of work involved in analysing the aberrations of a given system compares now even more favourably with that required in ray tracing methods, especially if skew rays are taken into account. (Notice that to trace rays through non-spherical surfaces an algebraic equation of degree n must be solved at each surface of the nth degree, and for every ray). Secondly, the computing scheme gives complete information concerning the contributions to the aberrations by the different surfaces(S§25(b)), And finally the adjustment of asphericities is a relatively easy matter; whereas with ray tracing methods alone the (finite) simultaneous adjustment at several surfaces of the asphericities would appear to be a most cumbersome.

It is hoped at a later date to supplement the development above with a systematisation of the means of calculating aberrations of order higher than the second. In particular it is hoped to give all expressions required for the calculation of the complete tertiary aberrations explicitly. Certain other problems may also be dealt with, such as the aberrations of optical systems with continuously varying refractive index, which will incidentally throw an interesting light on the mathematical foundations of the present algebraic methods.

The following table is to be regarded as a supplement to the table of §27. It contains the more important new symbols which occur in the text. The numbers in brackets again refer to the section in which the symbol first occurs.

Symbol	Meaning of symbol
$\underline{A}$	a/a' . (5)
a_i $(i=0,\dots,3)$	coefficient of $Y_1^i V_1^{3-i}$ in ε'_y , $(Z_1=W_1=0)$. (13)
b_i $(i=0,\dots,5)$	coefficient of $Y_1^i V_1^{5-i}$ in ε'_y , $(Z_1=W_1=0)$. (13)
$\underline{C}$	curvature of curve generating refracting surface. (2)
$\underline{\tilde{C}}$	$N(\underline{Y} - r\underline{V})$. (3)
$\underline{C}$	$N(\underline{Y} - r_0\underline{V})$. (3)
γ_i $(i=1,\dots,7)$	groups of terms in $\underline{J}$ and $\underline{L}$. (8)
D_i $(i=1,2)$	$D_1 = \sum_{i=1}^{3} a_i p^i$, $D_2 = \sum_{i=1}^{5} b_i p^i$. (13)
η	$\underline{Y}\cdot\underline{V} = YV + ZW$. (7)
ζ	$\underline{V}\cdot\underline{V} = V^2 + W^2$. (7)
θ_n	coefficient of χ^n in series expansion of $\underline{x}$. (2)

$\underline{i}$	unit normal to plane of incidence. (3)
$\underline{I}$	$\underline{C}/Nr_0$. (3)
$\underline{J}$	coefficient of $\underline{I}$ in $\Delta\underline{\Lambda}$. (4)
$\underline{L}$	coefficient of $\underline{Y}$ in $\Delta\underline{\Lambda}$. (4)
$\bar{N}$	ΔN . (4)
ξ	$\underline{Y}\cdot\underline{Y} = Y^2 + Z^2$. (7)
$\underline{P}$	$(r - x)/r$. (5)
p	paraxial location of entrance pupil . (13)
r	$\overrightarrow{AC}$. (2)
r_0	$\overrightarrow{AC_0}$. (2)
ρ	$\overrightarrow{PC}$. (2)
$\underline{R}$	$(1/r - 1/r_0)$. (9)
a_n	coefficient of χ^n in expansion of $\underline{C}$. (9)
$\underline{S}$	$r/\rho\sigma a'$. (4)
σ	radial polar coordinate used in specifying rays of a pencil with respect to pr.ray. (13)
$\bar{V}$	V_1 of principal ray. (13)
χ	$y^2 + z^2$. (2)
x	coordinate of points on surface . (2)
$\underline{X}$	$x\underline{S}$. (4)
ψ	angular coordinate corresponding to σ . (13)
y	coordinate of points on surface. (2)
$\bar{Y}$	Y_1 of principal ray. (13)
z	coordinate of points on surface . (2)

References.

Buchdahl, H. A.	1946.	Proc. Phys. Soc. LVIII. p. 545.
"	1947.	This thesis, p. 32 .
"	1948.	J. O. S. A. Vol. 38, No. 1. p. 14.
Conrady, A. E.	1929.	Applied Optics and Optical Design. O.U.P. Ch. VIII., p. 405.

THE ALGEBRAIC THEORY AND CALCULATION OF THE GEOMETRICAL HIGHER ORDER ABERRATIONS OF OPTICAL SYSTEMS.

III. Supplementary results concerning the aberrations of the general symmetrical optical system.

by H.A.Buchdahl, Department of Physics,
University of Tasmania.

Abstract: - In Part I of this supplement a systematization of the explicit algebraic developments arising from the author's previous work is carried out. The purpose of this is to make it possible to derive all expressions necessary for the computation of the aberrations to any required order according to a definite systematic procedure. In particular, all expressions necessary for the computation of the (exact) primary, secondary, and tertiary aberration coefficients of symmetrical optical systems are derived in detail by these means.

Part II deals with symmetrical optical systems in which the refractive index is a continuously variable function of position. This is of heuristic interest insofar as a "Lagrangian" treatment of the subject leads to a development which is in all respects analogous to the author's previous algebraic methods. Higher order aberrations are again obtainable by iteration.

In Part III a concise proof of the sine relation is given, first for the case of 'ordinary' symmetrical systems; and this is then extended to the case of the most general symmetrical system in which the medium may be both heterogeneous and anisotropic, the behaviour of systems of the type found in electron optics thereby being covered.

PART I: ON THE COMPUTATION OF TERTIARY AND HIGHER ORDER ABERRATIONS.

§1. Introduction.

In previous papers (Buchdahl, 1948(2), 1948(3); hereafter referred to by the letters S and A respectively) the author has dealt at some length with the general principles underlying his algebraic methods for the determination of the exact higher order aberrations of symmetrical optical systems. In particular, all series expansions necessary for the determination of the primary and secondary aberrations of such systems were given fully in explicit form in S §§20 and 21, and A §§7-10. On the basis of these, sample computations were actually carried out for the case of a Cooke Triplet (v. S §25 and A §12). It is apparent that it may be desirable in practice at times to compute the tertiary or even higher order aberration coefficients. We propose therefore to consider now a systematization of the means of obtaining the expansion of $\Delta\Lambda$ when going beyond secondary aberrations, as otherwise one runs the danger of becoming bogged in an unmanageable welter of different series and expressions. At the same time the recurrence of certain groups of terms is made more obvious, with consequent shortening of computing schemes used in actual calculations.

In particular the detailed expansion of $\Delta\Lambda$ will be extended to include the tertiary aberrations, and for the case of the latter the expressions corresponding to those of S §21 are also given in full.

§2. Fundamental sets of terms.

(a) The basic series of a given refracting surface is that which expresses the shape of its curve of intersection with a plane containing its axis of revolution, viz.

$$x = \sum_{n=1}^{\infty} \theta_n \chi^n \quad . \tag{2.1}$$

(The meaning of any symbol not defined here is as given in the tables of S§27 and A§14.) In virtue of the identity A(7.2), χ may be expressed as a function of ξ, η, ζ ,

$$\chi = \sum_{n=1}^{\infty} \chi^{(n)} \quad , \tag{2.2}$$

where as before $\chi^{(n)}$ is a homogeneous polynomial of degree n in ξ, η, ζ , the coefficients of which are now to be thought of as combinations of the a_n of A§2(b). Writing for later convenience $x^{(n)} = \beta_n$, the substitution of (2.2) in (2.1) yields

$$x = \sum_{n=1}^{\infty} \beta_n \quad . \tag{2.3}$$

In other words, the condition that the polynomials of degree s, ($s=1,2,\ldots$), in the expression

$$\sum_{n=1}^{\infty} \left\{ \theta_n \left[\xi - 2\eta \sum_{m=1}^{\infty} \beta_m + \zeta \left(\sum_{m=1}^{\infty} \beta_m\right)^2 \right]^n - \beta_n \right\} \tag{2.31}$$

vanish identically implies a recurrence relation for the β_m. Similarly we obtain, with $\underline{R}^{(n)} = \gamma_n$,

$$\underline{R} = \sum_{n=1}^{\infty} \gamma_n \quad . \tag{2.4}$$

Since $\underline{P} = (r - x)/r$ this incidentally gives, with $\underline{P}^{(n)} = p_n$,

$$\left.\begin{aligned} p_n &= a_0\beta_n - \sum_{m=1}^{n-1} \beta_{n-m}\gamma_m \quad , \qquad (n \geq 1) \\ p_0 &= 1 \quad . \end{aligned}\right\} \tag{2.5}$$

(b) In addition to the two fundamental sets of terms β_n and γ_n we introduce two further sets σ_n and τ_n , each of which is again a homogeneous polynomial of degree n in ξ, η, ζ . They are defined by

$$-\underline{s}_1/2(\underline{s}_1^{(0)}+\underline{s}_2^{(0)}) = -\frac{1}{2k}\{k^2(\alpha_0\eta-\zeta)+k^2\eta\underline{R}+(1+\zeta)\underline{P}\} = \sum_{n=0}^{\infty}\sigma_n\,, \tag{2.61}$$

$$-\underline{s}_2/2(\underline{s}_1^{(0)}+\underline{s}_2^{(0)}) = -\frac{1-k}{2k}\{k^2(\alpha_0^2\xi-2\alpha_0\eta+\zeta)+2k^2(\alpha_0\xi-\eta)\underline{R} + k^2\xi\underline{R}^2-(1+\zeta)\underline{P}^2\} = \sum_{n=0}^{\infty}\tau_n\,. \tag{2.62}$$

It is of importance to notice that the σ_n, τ_n, $(n \geq 2)$, are themselves expressed in terms of γ_m, p_m, that is, in effect, in terms of β_m, γ_m. In fact

$$\left.\begin{aligned}\sigma_0 &= -1/2k\\ \sigma_1 &= \sigma_0[k^2(\alpha_0\eta-\zeta)+p_1+\zeta]\\ \sigma_n &= \sigma_0(k^2\eta\gamma_{n-1}+p_n+\zeta p_{n-1}) \quad , (n\geq 2),\end{aligned}\right\} \tag{2.7}$$

$$\left.\begin{aligned}\tau_0 &= (1-k)/2k\\ \tau_1 &= -\tau_0[k^2(\alpha_0^2\xi-2\alpha_0\eta+\zeta)-2p_1-\zeta]\\ \tau_n &= -\tau_0[2k^2(\alpha_0\xi-\eta)\gamma_{n-1}+k^2\xi\sum_{m=1}^{n-2}\gamma_m\gamma_{n-m-1}\\ &\qquad -\sum_{m=0}^{n}p_mp_{n-m}-\zeta\sum_{m=0}^{n-1}p_mp_{n-m-1}]\,, \; (n\geq 2)\end{aligned}\right\} \tag{2.8}$$

(See also §6(b).) If in any term of a Σ the lower index exceeds the upper, that term is to be taken as zero.

§3. Solution of the equation for $\underline{S}$.

If the expressions (2.61) and (2.62) be now inserted into the quadratic equation $\underline{A}$(5.5) for $\underline{S}$, we have at once

$$\left(\sum_{n=0}^{\infty}\tau_n\right)\underline{S}^2 + 2\left(\sum_{n=0}^{\infty}\sigma_n\right)\underline{S} + \frac{1+k}{2k}(1+\zeta) = 0\,. \tag{3.1}$$

For the sake of convenience we write herafter S_n for $\underline{S}^{(n)}$. Also let

$$\Psi_n = \sum_{m=0}^{n}S_mS_{n-m}\quad . \tag{3.2}$$

Then (3.1) becomes

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(\tau_n\Psi_m+2\sigma_nS_m) + \frac{1+k}{2k}(1+\zeta) = 0\,. \tag{3.3}$$

Selecting the terms of degree n , we find

$$\left.\begin{aligned} S_0 &= 1 \\ S_1 &= \tau_1 + 2\sigma_1 + \frac{1+k}{2k}\varsigma \quad , \end{aligned}\right\} \quad (3.4)$$

and all subsequent S_n are obtained from the relation

$$\sum_{m=0}^{n} (\tau_{n-m}\Psi_m + 2\sigma_{n-m}S_m) \equiv 0 \quad , \quad (n \geq 2) . \qquad (3.5)$$

This equation may be put into a form in which it is more obviously in the nature of a recurrence relation for the S_n by separating out the terms $m=0$ and $m=n$. We thus get

$$S_n = \sum_{m=1}^{n-1}\{\tau_0 S_m S_{n-m} + \tau_{n-m}\sum_{r=1}^{m-1} S_r S_{m-r} + 2(\tau_{n-m} + \sigma_{n-m})S_m\} + (2\sigma_n + \tau_n) \quad . \qquad (3.6)$$

§4. Expression for $\Delta\Lambda$.

(a) In A§4 it was shown that $\underline{\Delta\Lambda}$ could be exhibited in the form

$$\underline{\Delta\Lambda} = \underline{J}\,\underline{I} + \underline{L}\,\underline{Y} \quad , \qquad (4.1)$$

where $\underline{J} = \bar{N}\{(\underline{S} - 1)y_0 - \underline{X}u_0'\}$

and $\underline{L} = \bar{N}\,\underline{R}\{\underline{S}y_0 - \underline{X}u_0'\}$, } (4.2)

with $\underline{X} = x\underline{S}$. (4.21)

Accordingly we have at once, since $\beta_0 = 0$,

$$\underline{X}^{(n)} = \sum_{m=0}^{n-1}\beta_{n-m}S_m \quad . \qquad (4.3)$$

Therefore $\underline{J}^{(n)} = \bar{N}\{y_0 S_n - u_0'\sum_{m=0}^{n-1}\beta_{n-m}S_m\}$, (4.4)

whence $\underline{L}^{(n)} = \bar{N}\; y_0\gamma_n + \sum_{m=1}^{n-1}\gamma_{n-m}\underline{J}^{(m)}$. (4.5)

(b) This essentially completes the systematisation of the method of expansion. We note how in proceeding from order $\underline{n-1}$ to order $\underline{n}$, β_n and γ_n , and thence σ_n and τ_n appear as essentially 'new' terms. But the equations (2.5), (2.7), (2.8), (3.6), and

(4.4), (4.5) clealy show how groups of terms occurring in the expressions for orders $< \underline{n}$ may be used in arriving at the terms of order $\underline{n}$. The simplicity of (4.5) also is a great advantage.

§5. β_n, γ_n, σ_n, τ_n, for $n = 1,2,3$.

(a) Straightforward calculation as in A§7 easily leads to the following terms, using

$$\chi = \xi - 2\theta_1\xi\eta + (-2\theta_2\xi^2\eta + \theta_1^2\xi^2\zeta + 4\theta_1^2\xi\eta^2) + O(8) \quad , \qquad (5.1)$$

$$\left.\begin{aligned}\gamma_1 &= \alpha_1\xi \\ \gamma_2 &= (\alpha_2 + \theta_1^2\alpha_1)\xi^2 - \alpha_0\alpha_1\xi\eta \\ \gamma_3 &= (\alpha_3 + \tfrac{1}{3}\alpha_0^2\alpha_2 + \theta_1\alpha_1^2 + 2\theta_1^4\alpha_1)\xi^3 - (2\alpha_0\alpha_2 + 6\theta_1^3\alpha_1 + \tfrac{1}{2}\alpha_1^2)\xi^2\eta \\ &\qquad + \theta_1^2\alpha_1\xi^2\zeta + \alpha_0^2\alpha_1\xi\eta^2 \quad ,\end{aligned}\right\}(5.2)$$

$$\left.\begin{aligned}\beta_1 &= \theta_1\xi \\ \beta_2 &= \theta_2\xi^2 - 2\theta_1^2\xi\eta \\ \beta_3 &= \theta_3\xi^3 - 6\theta_1\theta_2\xi^2\eta + \theta_1^3\xi^2\zeta + 4\theta_1^3\xi\eta^2 \quad .\end{aligned}\right\}(5.3)$$

[We are using the θ_n, α_n indifferently at this stage, according to convenience. They are related by ($\underline{v}$. $\underline{S}$§2(b))

$$\left.\begin{aligned}\theta_1 &= \tfrac{1}{2}\alpha_0 \\ \theta_2 &= \tfrac{1}{4}\alpha_1 + \theta_1^3 \\ \theta_3 &= \tfrac{1}{6}\alpha_2 + \theta_1^2\alpha_1 + 2\theta_1^5 \\ \theta_4 &= \tfrac{1}{8}(\alpha_3 + 6\theta_1^2\alpha_2 + 30\theta_1^4\alpha_1 + 3\theta_1\alpha_1^2 + 40\theta_1^5) \quad .\end{aligned}\right\}(5.4)$$]

(b) In virtue of (5.1) and (5.2) we at once obtain the p_n from (2.5),

$$\left.\begin{aligned}p_0 &= 1 \\ p_1 &= -2\theta_1^2\xi \\ p_2 &= -(\tfrac{3}{2}\theta_1\alpha_1 + 2\theta_1^4)\xi^2 + 4\theta_1^3\xi\eta \\ p_3 &= -(\tfrac{4}{3}\theta_1\alpha_2 + 4\theta_1^3\alpha_1 + \tfrac{1}{4}\alpha_1^2 + 4\theta_1^6)\xi^3 + (7\theta_1^2\alpha_1 + 12\theta_1^5)\xi^2\eta \\ &\qquad - 2\theta_1^4\xi^2\zeta - 8\theta_1^4\xi\eta^2 \quad .\end{aligned}\right\}(5.5)$$

196

This in turn leads to the second set of fundamental terms, in virtue of (2.7), (2.8), viz.

$$\left.\begin{aligned}
\sigma_0 &= -1/2k \\
\sigma_1 &= \tfrac{1}{2k}\left[\tfrac{1}{2}\alpha_0^2\xi - k^2\alpha_0\eta - (1-k^2)\zeta\right] \\
\sigma_2 &= \tfrac{1}{k}\left[(\tfrac{3}{4}\theta_1\alpha_1 + \theta_1^4)\xi^2 - (\tfrac{1}{2}k^2\alpha_1 + 2\theta_1^3)\xi\eta + \theta_1^2\xi\zeta\right] \\
\sigma_3 &= \tfrac{1}{2k}\Big[(\tfrac{4}{3}\theta_1\alpha_2 + 4\theta_1^3\alpha_1 + \tfrac{1}{4}\alpha_1^2 + 4\theta_1^6)\xi^3 - \left(k^2\alpha_2 + (k^2+7)\theta_1^2\alpha_1 + 12\theta_1^5\right)\xi^2\eta \\
&\quad + (\tfrac{3}{2}\theta_1\alpha_1 + 4\theta_1^4)\xi^2\zeta + (k^2\alpha_0\alpha_1 + 8\theta_1^4)\xi\eta^2 - 4\theta_1^3\xi\eta\zeta\Big],
\end{aligned}\right\} \quad (5.6)$$

$$\left.\begin{aligned}
\tau_0 &= (1-k)/2k \\
\tau_1 &= -\tfrac{1-k}{2k}\left[(1+k^2)\alpha_0^2\xi - 2k^2\alpha_0\eta - (1-k^2)\zeta\right] \\
\tau_2 &= -\tfrac{1-k}{2k}\left[(4k^2+3)\theta_1\alpha_1\xi^2 - (2k^2\alpha_1 + \alpha_0^3)\xi\eta + \alpha_0^2\xi\zeta\right] \\
\tau_3 &= -\tfrac{1-k}{2k}\Big[\left((2k^2+\tfrac{4}{3})\alpha_0\alpha_2 + (k^2+\tfrac{1}{2})\alpha_1^2 + (4k^2+2)\theta_1^3\alpha_1\right)\xi^3 \\
&\quad - \left(2k^2\alpha_2 + (10k^2+14)\theta_1^2\alpha_1 + 8\theta_1^5\right)\xi^2\eta + (3\theta_1\alpha_1 + 4\theta_1^4)\xi^2\zeta \\
&\quad + (2k^2\alpha_0\alpha_1 + \alpha_0^4)\xi\eta^2 - \alpha_0^3\xi\eta\zeta\Big] .
\end{aligned}\right\} \quad (5.7)$$

§6. On the computation of $\Delta\Delta^{(n)}$, $(n=1,2,3)$.

(a) From (3.6) we obtain for the S_n, $(n=1,2,3)$ the following expressions:

$$S_1 = \tau_1 + 2\sigma_1 + \frac{1+k}{2k}\zeta \qquad (6.11)$$

$$S_2 = \tau_0 S_1^2 + 2(\tau_1 + \sigma_1)S_1 + (\tau_2 + 2\sigma_2) \qquad (6.12)$$

$$S_3 = 2\tau_0 S_1 S_2 + 2(\tau_2+\sigma_2)S_1 + 2(\tau_1+\sigma_1)S_2 + \tau_1 S_1^2 + (\tau_3 + 2\sigma_3) . \qquad (6.13)$$

On the other hand, from (4.4), (4.5) we have

$$\left.\begin{aligned}
\underline{J}^{(1)} &= \bar{N}(y_0 S_1 - u_0'\beta_1) \\
\underline{L}^{(1)} &= \bar{N} y_0 \gamma_1 ,
\end{aligned}\right\} \quad (6.21)$$

$$\left.\begin{aligned}
\underline{J}^{(2)} &= \bar{N}\left[y_0 S_2 - u_0'(\beta_1 S_1 + \beta_2)\right] \\
\underline{L}^{(2)} &= \bar{N}\,\mathcal{Y}_2 + \gamma_1 \underline{J}^{(1)} ,
\end{aligned}\right\} \quad (6.22)$$

$$\left.\begin{aligned} \underline{J}^{(3)} &= \bar{N}[y_0 S_3 - u_0'(\beta_1 S_2 + \beta_2 S_1 + \beta_3)] \\ \underline{L}^{(3)} &= \bar{N} y_0 \gamma_3 + (\gamma_1 \underline{J}^{(2)} + \gamma_2 \underline{J}^{(1)}) \quad . \end{aligned}\right\} (6.23)$$

(b) It will be seen that, in general, we may proceed as follows. First the β_n, γ_n are computed to the required order, and from these the p_n. The σ_n are then easily determined from (2.7); i.e. for n=1,2,3

$$\left.\begin{aligned} \sigma_1 &= \sigma_0[k^2(\alpha_0\eta - \zeta) + p_1 + \zeta] \\ \sigma_2 &= \sigma_0[p_2 + \zeta p_1 + k^2\eta\gamma_1] \\ \sigma_3 &= \sigma_0[p_3 + \zeta p_2 + k^2\eta\gamma_2] \quad , \end{aligned}\right\} (6.3)$$

which requires only very few multiplications. We could then find the τ_n directly from (2.8). But it may be simpler to use the identity

$$\tau_n + 2(1-k)\sigma_n = -\tau_0\Big\{\xi\big[2k^2\alpha_0\gamma_{n-1} + 2\theta_1^2 p_{n-1} + k^2\sum_{m=1}^{n-1}\gamma_m\gamma_{n-m-1}\big] - \sum_{m=1}^{n-2} p_m(p_{n-m} + \xi p_{n-m-1})\Big\} \;, (n \geq 2), \quad (6.4)$$

which follows from (2.7) and (2.8), and the right hand side of which involves neither p_n nor γ_n. Thus, for n=2,3,

$$\left.\begin{aligned} \tau_2 + 2(1-k)\sigma_2 &= 4\tau_0(\theta_1^4 - k^2\theta_1\alpha_1)\xi^2 \\ \tau_3 + 2(1-k)\sigma_3 &= -\tau_0\Big\{\big[2k^2\alpha_0\alpha_2 + k^2\alpha_1^2 + (4k^2-6)\theta_1^3\alpha_1 - 8\theta_1^6\big]\xi \\ &\qquad - (2k^2\alpha_0^2\alpha_1 - 16\theta_1^5)\eta - 4\theta_1^4\zeta\Big\}\xi^2 \quad . \end{aligned}\right\} (6.5)$$

The S_n are then found from (6.13); which task may be somewhat shortened by observing that in S_3 the combinations S_1^2 and $2(\tau_1 + \sigma_1)$ for instance occur already in S_2. The determination of $\underline{J}^{(n)}$ and $\underline{L}^{(n)}$ (n=1,2,3) is now not a very lengthy matter. The $\underline{L}^{(n)}$ especially involve little further calculation.

(c) The forms of $\Delta\underline{\Lambda}^{(n)}$, (n=1,2), may not be quite as convenient as those of A§9. But when tertiary (or higher) aberrations have to be computed the position is rather different. It is of course possible that a lengthy search might again reveal combinations of terms which could make actual calculation somewhat shorter if n were arbitrarily restricted to be, say, < 4. But it is doubtful whether much labour would be saved in the end, and certainly the great advantage of leaving open the possibility of systematically proceeding to some higher order should the need to do so subsequently arise in practice would be lost. In any case, with computing schemes suitably arranged the labour involved in finding $\Delta\Lambda^{(n)}$, at any rate for $n < 4$, is such as to make routine computations of this kind quite feasible in practice. The prototype of such schemes may, in a general way, be taken as that of S§25(b).

§7. Spherical surfaces.

Spherical surfaces are of particular interest partly because of their very frequent occurrence in practice, and partly because it may be desirable to keep the "spherical terms", (i.e. the terms of $\Delta\Lambda$ which survive when we set $\alpha_i = 0$, $(i \geq 1)$), separate in the course of computation, as was done in A§9. Putting $\alpha_i = 0$, $(i \geq 1)$, great simplifications naturally arise. Thus the β_n are now given by the series solution of the equation

$$(1 + \zeta)x^2 - 2(r_0 + \eta)x + \xi = 0 \quad , \tag{7.1}$$

which yields the recurrence formula for the β_n, viz.

$$\beta_n = \theta_1\left[(\theta_1\xi - 2\eta)\beta_{n-1} + \sum_{m=1}^{n-2}\beta_m(\beta_{n-m} + \zeta\beta_{n-m-1})\right] \quad . \tag{7.2}$$

All the γ_n on the other hand vanish identically, whence

$$p_n = -\alpha_o\beta_n \quad . \tag{7.3}$$

The $\sigma_o, \tau_o, \sigma_1, \tau_1$ are unchanged; but for the σ_2, σ_3 we have simply

$$\left.\begin{aligned} \sigma_2 &= \theta_1^2\xi(\theta_1^2\xi - 2\theta_1\eta + \zeta)/k \\ \sigma_3 &= 2\theta_1(\theta_1\xi - \eta)\sigma_2 \quad ; \end{aligned}\right\} \tag{7.4}$$

whilst the τ_2, τ_3 now follow most easily from (5.7), i.e.

$$\left.\begin{aligned} \tau_2 &= \tau_o\alpha_o^2\xi(\alpha_o\eta - \zeta) \\ \tau_3 &= \theta_1(\theta_1\xi + 2\eta)\tau_2 \quad . \end{aligned}\right\} \tag{7.5}$$

The simplicity of these formulae is remarkable, and the calculation of the S_n and thence of $\underline{J}^{(n)}$ is particularly easy. $\underline{L}^{(n)}$ of course vanishes identically.

§8. Tertiary increments in explicit form.

(a) The problem of iteration was fully discussed in S§5(a),(b). For the purpose of setting out a practical computing scheme it is convenient to have available the coefficients $g^{(n)}_{\mu\nu}$, $\bar{g}^{(n)}_{\mu\nu}$ in the expansion S(5.3) of $\Delta\Lambda$ on the one hand in terms $\underline{g}^{(s)}_{\mu\nu|p}$, $\underline{\bar{g}}^{(s)}_{\mu\nu|p}$ and $\underline{g}^{(s)}_{\mu\nu|q}$, $\underline{\bar{g}}^{(s)}_{\mu\nu|q}$, $(s=1,2,\ldots,m)$, (the determination of which we have considered above); and on the other hand the sums which arise from them in the course of iteration, viz. $G^{(s)}_{\mu\nu|p}, \ldots$, where $s \leq m-1$, (v. S(3.101)). In S§21 we gave fully the expressions for the $g^{(2)}_{\mu\nu}$, $\bar{g}^{(2)}_{\mu\nu}$, (which split up into their respective p and q 'components'). We now give, in (b), the corresponding expressions for the $g^{(3)}_{\mu\nu}$, $\bar{g}^{(3)}_{\mu\nu}$ in full. They are obtained by substituting

$$\left.\begin{aligned} &\underline{Y}_1 - (A_q\underline{Y}_1\xi_1 + \ldots + \bar{C}_q\underline{V}_1\zeta_1) - (S_{1q}\underline{Y}_1\xi_1^2 + \ldots + \bar{S}_{6q}\underline{V}_1\zeta_1^2) \\ &\underline{V}_1 + (A_p\underline{Y}_1\xi_1 + \ldots + \bar{C}_p\underline{V}_1\zeta_1) + (S_{1p}\underline{Y}_1\xi_1^2 + \ldots + \bar{S}_{6p}\underline{V}_1\zeta_1^2) \end{aligned}\right\} \tag{8.1}$$

for $\underline{Y}_1$, $\underline{V}_1$ respectively in the series

$$\underline{\Delta\Lambda} = (\underline{a}Y_1\xi_1+\ldots+\underline{\bar{c}}V_1\zeta_1) + (\underline{s}_1Y_1\xi_1^2+\ldots\ \underline{\bar{s}}_6V_1\zeta_1^2) + (\underline{t}_1Y_1\xi_1^3+\ldots+\underline{\bar{t}}_{10}V_1\zeta_1^3), \qquad (8.2)$$

using the notation of $\underline{S}$(3.4),(3.10). Selecting the coefficients of the terms of the seventh degree then gives rise to the expressions set out below. The corresponding expressions for the $g^{(n)}_{\mu\nu}$, $\bar{g}^{(n)}_{\mu\nu}$, ($m \geqslant 4$) may be arrived at in a precisely similar manner.

As before each equation splits up into two components. So that if we are interested in the tertiary aberrations for arbitrary object position the increments, considered collectively, consist of the sum of 2·692 = 1384 separate terms. But the following very great simplifications should be considered:

(i) the $(\bar{t})_{ip}$ once being determined, the $(\bar{t})_{iq}$ follow very easily by means of relatively few multiplications.

(ii) Most of the multiplications are the product of primary coefficients, two at a time. There are only 78 such products, so that many of them must recur several times over. (Also, these products are rapidly computed in groups of ~~twelve~~ 12,11,10,..., leaving one factor on the calculating machine during each calculation of one such group).

(iii) whole groups of terms recur in different coefficients, e.g. consider the brackets containing S_{2p}.

(b) It should be noticed that if $N_1 \neq 1$ every capital roman symbol below is to be thought of as divided by N_1.

Coeffs. of

$Y\xi^3 = t_1$

$$\underline{t}_1 - 5A_q\,\underline{s}_1 + A_p\,\underline{\bar{s}}_1 + A_p\,\underline{s}_2 + (-3S_{1q} + 3A_q^{\,2})\underline{a} + (S_{1p} - 2A_p\,A_q)\underline{\bar{a}} + (S_{1p} - 2A_p\,A_q)\underline{b} + A_p^{\,2}\,\underline{\bar{b}} + A_p^{\,2}\,\underline{c}$$

$V\xi^3 = \bar{t}_1$

$$\underline{\bar{t}}_1 - \bar{A}_q\,\underline{s}_1 + (\bar{A}_p - 4A_q)\underline{\bar{s}}_1 + A_p\,\underline{\bar{s}}_2 + (-\bar{S}_{1q} + 2A_q\,\bar{A}_q)\underline{a} + (\bar{S}_{1p} - 2S_{1q} + A_q^{\,2} - 2\bar{A}_p\,A_q)\underline{a} + (-A_p\,\bar{A}_q)\underline{b} + (S_{1p} + A_p\,\bar{A}_p - A_p\,A_q)\underline{\bar{b}} + A_p^{\,2}\,\underline{\bar{c}}$$

$Y\xi^2\eta = t_2$

$$\begin{aligned}\underline{t}_2 &- (4\bar{A}_q + 5B_q)\underline{s}_1 + B_p\,\underline{\bar{s}}_1 + (\bar{A}_p - 4A_q + B_p)\underline{s}_2 + A_p\,\underline{\bar{s}}_2 + 2A_p\,\underline{s}_3 + 2A_p\,\underline{s}_4\\ &+ (-2\bar{S}_{1q} - 3S_{2q} + 4A_q\,\bar{A}_q + 6A_q\,B_q)\underline{a} + (S_{2p} - 2A_p\,\bar{A}_q - 2A_p\,B_q - 2A_q\,B_p)\underline{\bar{a}}\\ &+ (\bar{S}_{1p} + S_{2p} - 2S_{1q} - A_p\,\bar{A}_q + A_q^{\,2} - 2\bar{A}_p\,A_q - 2A_p\,B_q - 2A_q\,B_p)\underline{b}\\ &+ (S_{1p} - A_p\,A_q + A_p\,A_p + 2A_p\,B_p)\underline{\bar{b}} + (2S_{1p} + 2A_p\,\bar{A}_p - 2A_p\,A_q + 2A_p\,B_p)\underline{c} + 2A_p^{\,2}\,\underline{\bar{c}}\end{aligned}$$

$V\xi^2\eta = \bar{t}_2$

$$\begin{aligned}\underline{\bar{t}}_2 &- B_q\,\underline{s}_1 + (B_p - 4\bar{A}_q - 4B_q)\underline{\bar{s}}_1 - \bar{A}_q\,\underline{s}_2 + (2\bar{A}_p - 3A_q + B_p)\underline{\bar{s}}_2 + 2A_p\,\underline{s}_3 + 2A_p\,\underline{s}_4\\ &+ (-\bar{S}_{2q} + 2\bar{A}_q^{\,2} + 2A_q\,\bar{B}_q + 2\bar{A}_q\,B_q)\underline{a}\\ &+ (-2\bar{S}_{1q} - 2S_{2q} + \bar{S}_{2p} - 2\bar{A}_p\,\bar{A}_q + 2A_q\,\bar{A}_q - 2A_q\,\bar{B}_p + 2A_q\,B_q - 2\bar{A}_p\,B_q)\underline{\bar{a}}\\ &+ (-S_{1q} - \bar{A}_p\,\bar{A}_q + A_q\,\bar{A}_q - A_p\,\bar{B}_q - \bar{A}_q\,B_p)\underline{b}\\ &+ (2\bar{S}_{1p} + S_{2p} - S_{1q} - A_p\,A_q + A_p\,\bar{B}_p - A_p\,B_q - A_q\,B_p + \bar{A}_p^{\,2} - 2A_q\,\bar{A}_p + \bar{A}_p\,B_p)\underline{\bar{b}}\\ &+ (-2A_p\,\bar{A}_q)\underline{c} + (2S_{1p} + 4A_p\,\bar{A}_p + 2A_p\,B_p)\underline{\bar{c}}\end{aligned}$$

$Y\xi^2\zeta = t_3$

$$\begin{aligned}\underline{t}_3 &- 5C_q\,\underline{s}_1 + C_p\,\underline{\bar{s}}_1 + (C_p - \bar{A}_q)\underline{s}_2 + (2\bar{A}_p - 3A_q)\underline{s}_3 + A_p\,\underline{\bar{s}}_3 + A_p\,\underline{s}_5\\ &+ (-3S_{3q} + \bar{A}_q^{\,2} + 6A_q\,C_q)\underline{a} + (S_{3p} - 2A_q\,C_p - 2A_p\,C_q)\underline{\bar{a}}\\ &+ (S_{3p} - \bar{S}_{1q} + A_q\,\bar{A}_q - \bar{A}_p\,\bar{A}_q - 2A_p\,C_q - 2A_q\,C_p)\underline{b} + (2A_p\,C_p - A_p\,\bar{A}_q)\underline{\bar{b}}\\ &+ (2\bar{S}_{1p} - S_{1q} + \bar{A}_p^{\,2} + 2A_p\,C_p - 2A_q\,\bar{A}_p)\underline{c} + (S_{1p} + 2A_p\,\bar{A}_p)\underline{\bar{c}}\end{aligned}$$

2a

202

Coeffs. of

$V\xi^2\zeta$ $\quad \bar t_3 - \bar C_q s_1 + (\bar C_p - 4C_q)\bar s_1 + (C_p - \bar A_q)\bar s_2 - \bar A_q s_3 + (3\bar A_p - 2A_q)\bar s_3 + A_p \bar s_5$

$\quad + (-\bar S_{3q} + 2A_q \bar C_q + 2\bar A_q C_q)a + (-2\bar S_{3q} + \bar S_{3p} + \bar A_q^2 + 2A_q C_q - 2A_q \bar C_p - 2\bar A_p C_q)\bar a$

$= \bar t_3$ $\quad + (\bar A_q^2 - \bar A_q C_p - A_p \bar C_q)b + (S_{3p} - \bar S_{1q} - A_p C_q + A_p \bar C_p - 2\bar A_p \bar A_q - A_q C_p + \bar A_p C_p)\bar b$

$\quad + (-\bar S_{1q} - 2\bar A_p A_q)c + (3\bar S_{1p} + 3\bar A_p^2 + 2A_p C_p)\bar c$

$V\xi\eta^2$ $\quad t_4 - 4\bar B_q s_1 + (\bar B_p - 4B_q - 2\bar A_q)s_2 + B_p \bar s_2 + 2B_p s_3 + (2A_p + 2B_p - 3A_q)s_4 + A_p \bar s_4$

$\quad + 2A_p s_5 + (-2\bar S_{2q} - 3S_{4q} + 4A_q \bar B_q + 4\bar A_q B_q + 3B_q^2)a$

$\quad + (S_{4p} - 2A_p \bar B_q - 2\bar A_q B_p - 2B_p B_q)\bar a$

$= t_4$ $\quad + (\bar S_{2p} + S_{4p} - 2S_{2q} - A_p \bar B_q + 2A_q B_q - 2A_q \bar B_p - 2\bar A_p B_q - \bar A_q B_p - 2B_p B_q)b$

$\quad + (S_{2p} + A_p \bar B_p - A_p B_q + \bar A_p B_p - A_q B_p + B_p^2)\bar b$

$\quad + (2S_{2p} + 2A_p \bar B_p - 2A_p B_q + 2\bar A_p B_p - 2A_q B_p + B_p^2)c + 4A_p B_p \bar c$

$V\xi\eta^2$ $\quad \bar t_4 - 4\bar B_q \bar s_1 - \bar B_q s_2 + (2\bar B_p - 3B_q - 2\bar A_q)\bar s_2 + 2B_p \bar s_3 - \bar A_q s_4 + (3\bar A_p - 2A_q + 2B_p)\bar s_4$

$\quad + 2A_p \bar s_5 + (-\bar S_{4q} + 4\bar A_q \bar B_q + 2B_q \bar B_q)a$

$\quad + (-2\bar S_{2q} - 2S_{4q} + \bar S_{4p} + B_q^2 + 2A_q \bar B_q - 2\bar A_p \bar B_q + 2\bar A_q B_q - 2\bar A_q \bar B_p - 2\bar B_p B_q)\bar a$

$= \bar t_4$ $\quad + (-\bar S_{2q} + A_q \bar B_q - \bar A_p \bar B_q + \bar A_q B_q - \bar A_q \bar B_p - B_p \bar B_q)b$

$\quad + (2\bar S_{2p} + S_{4p} - S_{2q} - A_p \bar B_q - 2A_q \bar B_p + 2\bar A_p \bar B_p + B_p \bar B_p - B_p B_q - 2\bar A_p B_q - \bar A_q B_p)\bar b$

$\quad + (-2\bar A_q B_p - 2A_p \bar B_q)c + (2S_{2p} + 4\bar A_p B_p + 4A_p \bar B_p + B_p^2)\bar c$

Coeff. of

$$
\begin{aligned}
Y\xi\eta\zeta \qquad & \underline{t}_5 - 4\bar{C}_q\,\underline{s}_1 + (\bar{C}_p - 4C_q - \bar{B}_q)\underline{s}_2 + C_p\,\underline{\bar{s}}_2 + (2\bar{B}_p - 2\bar{A}_q - 3B_q + 2C_p)\underline{s}_3 + B_p\,\underline{\bar{s}}_3 \\
& + (2C_p - 2\bar{A}_q)\underline{s}_4 + (3\bar{A}_p - 2A_q + B_p)\underline{s}_5 + A_p\,\underline{\bar{s}}_5 + 4A_p\,\underline{\bar{s}}_6 \\
& + (-2\bar{S}_{3q} - 3S_{5q} + 4A_q\bar{C}_q + 2\bar{A}_q\bar{B}_q + 4\bar{A}_q C_q + 6B_q C_q)\underline{a} \\
& + (S_{5p} - 2A_p\bar{C}_q - 2\bar{A}_q C_p - 2B_p C_q - 2B_q C_p)\underline{\bar{a}} \\
= t_5 \qquad & + (\bar{S}_{3p} + S_{5p} - \bar{S}_{2q} - 2S_{3q} - A_p\bar{C}_q + A_q\bar{B}_q - \bar{A}_p\bar{B}_q + 2A_qC_q - 2\bar{A}_pC_q - 2A_q\bar{C}_p + \bar{A}_qB_q - \bar{A}_q\bar{B}_p - \bar{A}_qC_p - 2B_pC_q - 2B_qC_p)\underline{b} \\
& + (S_{3p} + A_p\bar{C}_p - A_pC_q - A_p\bar{B}_q - \bar{A}_qB_p + \bar{A}_pC_p - A_qC_p + 2B_pC_p)\underline{\bar{b}} \\
& + (2\bar{S}_{2p} + 2S_{3p} - S_{2q} + 2A_p\bar{C}_p - 2A_pC_q + 2\bar{A}_p\bar{B}_p - 2\bar{A}_pB_q + 2\bar{A}_pC_p - 2A_qC_p + 2B_pC_p - 2A_q\bar{B}_p)\underline{c} \\
& + (S_{2p} + 2A_p\bar{B}_p + 2\bar{A}_pB_p + 4A_pC_p)\underline{\bar{c}}
\end{aligned}
$$

$$
\begin{aligned}
V\xi\eta\zeta \qquad & \underline{t}_5 - 4\bar{C}_q\,\underline{\bar{s}}_1 - \bar{C}_q\,\underline{s}_2 + (2\bar{C}_p - 3C_q - \bar{B}_q)\underline{\bar{s}}_2 + (-\bar{B}_q)\underline{s}_3 + (3\bar{B}_p - 2\bar{A}_q - 2B_q + 2C_p)\underline{\bar{s}}_3 \\
& + (2C_p - 2\bar{A}_q)\underline{s}_4 + (-\bar{A}_q)\underline{s}_5 + (4\bar{A}_p + B_p - A_q)\underline{\bar{s}}_5 + 4A_p\,\underline{\bar{s}}_6 \\
& + (-\bar{S}_{5q} + 4\bar{A}_q\bar{C}_q + 2B_q\bar{C}_q + 2\bar{B}_qC_q)\underline{a} \\
& + (-2\bar{S}_{3q} - 2S_{5q} + \bar{S}_{5p} + 2A_q\bar{C}_q - 2\bar{A}_p\bar{C}_q + 2\bar{A}_q\bar{B}_q + 2\bar{A}_qC_q - 2\bar{A}_q\bar{C}_p + 2B_qC_q - 2\bar{B}_pC_q - 2B_q\bar{C}_p)\underline{a} \\
= t_5 \qquad & + (-\bar{S}_{3q} - \bar{A}_p\bar{C}_q + A_q\bar{C}_q + 2\bar{A}_q\bar{B}_q + \bar{A}_qC_q - \bar{A}_q\bar{C}_p - B_p\bar{C}_q - \bar{B}_qC_p)\underline{b} \\
& + (2\bar{S}_{3p} + S_{5p} - \bar{S}_{2q} - S_{3q} - A_p\bar{C}_q - 2\bar{A}_p\bar{B}_q - 2\bar{A}_pC_q + 2\bar{A}_p\bar{C}_p - 2\bar{A}_q\bar{B}_p - \bar{A}_qC_p - 2A_q\bar{C}_p + B_p\bar{C}_p - B_pC_q - B_qC_p + \bar{B}_pC_p)\underline{b} \\
& + (-S_{2q} - 2\bar{A}_qB_p - 2\bar{A}_qC_p - 2\bar{A}_p\bar{B}_q - 2A_p\bar{C}_q)\underline{c} \\
& + (3S_{2p} + 2S_{3p} + 4A_p\bar{C}_p + 6\bar{A}_p\bar{B}_p + 4\bar{A}_pC_p + 2B_pC_p)\underline{c}
\end{aligned}
$$

$$
\begin{aligned}
Y\xi\zeta^2 \qquad & \underline{t}_6 - \bar{C}_q\,\underline{s}_2 + (2\bar{C}_p - 3C_q)\underline{s}_3 + C_p\,\underline{\bar{s}}_3 + (C_p - \bar{A}_q)\underline{s}_5 + (4\bar{A}_p - A_q)\underline{s}_6 + A_p\,\underline{\bar{s}}_6 \\
& + (-3S_{6q} + 2\bar{A}_q\bar{C}_q + 3S_q^2)\underline{a} + (S_{6p} - 2C_pC_q)\underline{\bar{a}} \\
= t_6 \qquad & + (S_{6p} - \bar{S}_{3q} - \bar{A}_p\bar{C}_q + A_q\bar{C}_q + \bar{A}_qC_q - \bar{A}_q\bar{C}_p - 2C_pC_q)\underline{b} + (-A_p\bar{C}_q - \bar{A}_qC_p + C_p^2)\underline{\bar{b}} \\
& + (2\bar{S}_{3p} - S_{3q} + 2\bar{A}_p\bar{C}_p - 2\bar{A}_pC_q - 2A_q\bar{C}_p + C_p^2)\underline{c} + (S_{3p} + 2A_p\bar{C}_p + 2\bar{A}_pC_p)\underline{\bar{c}}
\end{aligned}
$$

Coeff. of

$V\,\xi\zeta^2$ $= \bar{t}_6$

$$\begin{aligned}
&\underline{\bar{t}}_6 - \bar{C}_q \underline{\bar{s}}_2 - \bar{C}_q \underline{s}_3 + (3\bar{C}_p - 2C_q)\underline{\bar{s}}_3 + (C_p - \bar{A}_q)\underline{\bar{s}}_5 - \bar{A}_q \underline{s}_6 + 5\bar{A}_p \underline{\bar{s}}_6 + (-S_{6q} + 2C_q C_q)\underline{a} \\
&+ (-2S_{6q} + \bar{S}_{6p} + 2\bar{A}_q \bar{C}_q - 2\bar{C}_p C_q + C_q^2)\underline{\bar{a}} + (2\bar{A}_q \bar{C}_q - C_p \bar{C}_q)\underline{b} \\
&+ (S_{6p} - \bar{S}_{3q} - 2\bar{A}_p \bar{C}_q - 2\bar{A}_q \bar{C}_p + C_p \bar{C}_p - C_p C_q)\underline{\bar{b}} + (-\bar{S}_{3q} - 2\bar{A}_q C_p - 2\bar{A}_p C_q)\underline{c} \\
&+ (3\bar{S}_{3p} + 6\bar{A}_p \bar{C}_p + C_p^2)\underline{\bar{c}}
\end{aligned}$$

$Y\,\eta^3$ $= t_7$

$$\begin{aligned}
&\underline{t}_7 - 2\bar{B}_q \underline{s}_2 + (2\bar{B}_p - 3B_q)\underline{s}_4 + B_p \underline{\bar{s}}_4 + 2B_p \underline{s}_5 + (-2\bar{S}_{4q} + 4B_q \bar{B}_q)\underline{a} + (-2B_q \bar{B}_q)\underline{\bar{a}} \\
&+ (\bar{S}_{4p} - 2S_{4q} - B_p \bar{B}_q + B_q^2 - 2\bar{B}_p B_q)\underline{b} + (S_{4p} + B_p \bar{B}_p - B_p B_q)\underline{\bar{b}} \\
&+ (2S_{4p} + 2B_p \bar{B}_p - 2B_p B_q)\underline{c} + 2B_p^2 \underline{\bar{c}}
\end{aligned}$$

$V\,\eta^3$ $= \bar{t}_7$

$$\begin{aligned}
&\underline{\bar{t}}_7 - 2\bar{B}_q \underline{\bar{s}}_2 - \bar{B}_q \underline{s}_4 + (3\bar{B}_p - 2B_q)\underline{\bar{s}}_4 + 2B_p \underline{\bar{s}}_5 + 2\bar{B}_q^2 \underline{a} + (-2\bar{S}_{4q} + 2B_q \bar{B}_q - 2B_p \bar{B}_q)\underline{\bar{a}} \\
&+ (-\bar{S}_{4q} - \bar{B}_p \bar{B}_q + B_q \bar{B}_q)\underline{b} + (2\bar{S}_{4p} - S_{4q} + \bar{B}_p^2 - 2\bar{B}_p B_q - B_p \bar{B}_q)\underline{\bar{b}} + (-2B_p \bar{B}_q)\underline{c} \\
&+ (2S_{4p} + 4B_p \bar{B}_p)\underline{\bar{c}}
\end{aligned}$$

Coeff. of

$$
\begin{aligned}
Y\eta^2\zeta \qquad & \underline{t}_8 - 2\bar C_q\,\underline{s}_2 - 2\bar B_q\,\underline{s}_3 + (2\bar C_p - 3C_q - 2\bar B_q)\underline{s}_4 + C_p\,\underline{\bar s}_4 + (3\bar B_p - 2B_q + 2C_p)\underline{s}_5 + B_p\,\underline{\bar s}_5 \\
& + 4B_p\,\underline{s}_6 + (-2\bar S_{5q} + \bar B_q^2 + 4B_q\bar C_q + 4\bar B_q C_q)\underline{a} + (-2\bar B_q C_p - 2B_p\bar C_q)\underline{\bar a} \\
& + (\bar S_{5p} - \bar S_{4q} - 2S_{5q} - B_p\bar C_q + B_q\bar B_q + 2B_q C_q - 2B_q\bar C_p - \bar B_p\bar B_q - 2\bar B_p C_q - \bar B_q C_p)\underline{b} \\
= t_8 \qquad & + (S_{5p} + B_p\bar C_p - B_p C_q + \bar B_p C_p - B_q C_p - B_p\bar B_q)\underline{\bar b} \\
& + (2\bar S_{4p} + 2S_{5p} - S_{4q} + \bar B_p^2 + 2B_p\bar C_p - 2B_p C_q + 2\bar B_p C_p - 2B_q C_p - 2B_q\bar B_p)\underline{c} \\
& + (S_{4p} + 2B_p\bar B_p + 4B_p C_p)\underline{\bar c}
\end{aligned}
$$

$$
\begin{aligned}
V\eta^2\zeta \qquad & \underline{\bar t}_8 - 2\bar C_q\,\underline{\bar s}_2 - 2\bar B_q\,\underline{s}_3 - \bar C_q\,\underline{s}_4 + (3\bar C_p - 2C_q - 2\bar B_q)\underline{\bar s}_4 - \bar B_q\,\underline{s}_5 + (4\bar B_p - B_q + 2C_p)\underline{\bar s}_5 \\
& + 4B_p\,\underline{\bar s}_6 + 4\bar B_q\bar C_q\,\underline{a} + (-2\bar S_{5q} + \bar B_q^2 + 2B_q\bar C_q - 2\bar B_p\bar C_q + 2\bar B_q C_q - 2\bar B_q\bar C_p)\underline{\bar a} \\
= \bar t_8 \qquad & + (-\bar S_{5q} + B_q\bar C_q - \bar B_p\bar C_q + B_q^2 - \bar B_q\bar C_p + \bar B_q C_q)\underline{b} \\
& + (2\bar S_{5p} - \bar S_{4q} - S_{5q} + 2\bar B_p\bar C_p - 2\bar B_p C_q - B_p\bar C_q - 2\bar B_p\bar B_q - 2B_q\bar C_p - \bar B_q C_p)\underline{\bar b} \\
& + (-\bar S_{4q} - 2\bar B_q\bar B_p - 2\bar B_q C_p - 2B_p\bar C_q)\underline{c} + (3\bar S_{4p} + 2S_{5p} + 3\bar B_p^2 + 4B_p\bar C_p + 4\bar B_p C_p)\underline{\bar c}
\end{aligned}
$$

$$
\begin{aligned}
Y\eta\zeta^2 \qquad & \underline{t}_9 - 2\bar C_q\,\underline{s}_3 - 2\bar C_q\,\underline{s}_4 + (3\bar C_p - 2C_q - \bar B_q)\underline{s}_5 + C_p\,\underline{\bar s}_5 + (4\bar B_p - B_q + 4C_p)\underline{s}_6 + B_p\,\underline{\bar s}_6 \\
& + (-2\bar S_{6q} + 2\bar B_q\bar C_q + 4C_q\bar C_q)\underline{a} - 2C_p\bar C_q\,\underline{\bar a} \\
& + (\bar S_{6p} - \bar S_{5q} - 2S_{6q} + B_q\bar C_q - \bar B_p\bar C_q - \bar B_q\bar C_p + \bar B_q C_q - C_p\bar C_q - 2\bar C_p C_q + C_q^2)\underline{b} \\
= t_9 \qquad & + (S_{6p} - B_p\bar C_q - \bar B_q C_p + C_p\bar C_p - C_p C_q)\underline{\bar b} \\
& + (2\bar S_{5p} + 2S_{6p} - S_{5q} + 2\bar B_p\bar C_p - 2\bar B_p C_q + 2C_p\bar C_p - 2C_p C_q - 2B_q\bar C_p)\underline{c} \\
& + (S_{5p} + 2B_p\bar C_p + 2\bar B_p C_p + 2C_p^2)\underline{\bar c}
\end{aligned}
$$

Coeff. of

$V \eta \zeta^2$ $\quad \underline{\bar{t}}_9 - 2\bar{C}_q \underline{\bar{s}}_3 - 2\bar{C}_q \underline{\bar{s}}_4 - \bar{C}_q \underline{s}_5 + (4\bar{C}_p - \bar{B}_q - C_q)\underline{\bar{s}}_5 - \bar{B}_q \underline{s}_6 + (5\bar{B}_p + 4C_p)\underline{\bar{s}}_6 + (2\bar{C}_q^{\,2})\underline{a}$

$\quad + (-2\bar{S}_{6q} + 2\bar{B}_q \bar{C}_q + 2C_q \bar{C}_q - 2\bar{C}_p \bar{C}_q)\underline{\bar{a}} + (-\bar{S}_{6q} + 2\bar{B}_q \bar{C}_q + C_q \bar{C}_q - \bar{C}_p C_q)\underline{b}$

$= \bar{t}_9$ $\quad + (2\bar{S}_{6p} - \bar{S}_{5q} - S_{6q} - 2\bar{B}_p \bar{C}_q - 2\bar{B}_q \bar{C}_p + \bar{C}_p^{\,2} - 2\bar{C}_p C_q - C_p \bar{C}_q)\underline{\bar{b}}$

$\quad + (-\bar{S}_{5q} - 2\bar{B}_q \bar{C}_p - 2\bar{B}_p \bar{C}_q - 2C_p \bar{C}_q)\underline{c} + (3\bar{S}_{5p} + 2S_{6p} + 6\bar{B}_p \bar{C}_p + 4C_p \bar{C}_p)\underline{\bar{c}}$

$Y \zeta^3$ $\quad \underline{t}_{10} - \bar{C}_q \underline{s}_5 + (4\bar{C}_p - C_q)\underline{s}_6 + C_p \underline{\bar{s}}_6 + (\bar{C}_q^{\,2})\underline{a} + (-\bar{S}_{6q} + C_q \bar{C}_q - \bar{C}_p \bar{C}_q)\underline{b}$

$= t_{10}$ $\quad + (-C_p \bar{C}_q)\underline{\bar{b}} + (2\bar{S}_{6p} - S_{6q} - 2C_q \bar{C}_p + \bar{C}_p^{\,2})\underline{c} + (S_{6p} + 2C_p \bar{C}_p)\underline{\bar{c}}$

$V \zeta^3$ $\quad \underline{\bar{t}}_{10} - \bar{C}_q \underline{\bar{s}}_5 - \bar{C}_q \underline{s}_6 + 5\bar{C}_p \underline{\bar{s}}_6 + (\bar{C}_q^{\,2})\underline{\bar{a}} + (\bar{C}_q^{\,2})\underline{b} + (-\bar{S}_{6q} - 2\bar{C}_p \bar{C}_q)\underline{\bar{b}}$

$= \bar{t}_{10}$ $\quad + (-\bar{S}_{6q} - 2\bar{C}_p C_q)\underline{c} + (3\bar{S}_{6p} + 3\bar{C}_p^{\,2})\underline{\bar{c}}$

§9. Conclusion.

In conclusion, rather than repeat the remarks of S§27 and A§14 we take the opportunity of referring to some recent remarks of T. Smith, (Smith, 1948). It is reasonable to claim that the methods set out above provide just the algebraic tool for the examination of higher order aberrations which Mr. Smith wants to see developed, especially as it provides an easy method for the control of aberrations by means of manipulation of the asphericities of refracting surfaces constituting the optical system.

In complete accordance with his remarks it may be worth while to examine possible combinations of frequently recurring terms, as was already pointed out in §6(c) above, coupled possibly with the construction of tables of certain functions of k, etc.

Finally it may be mentioned that as long as the present canonical coordinates are used it is not likely that any algebraic method can be essentially shorter than this one. These particular coordinates were singled out by the author from amongst other linear coordinates, (v. S§§12 and 17), bacause they seemed especially convenient in practical work, and because they are always allowed coordinates, whatever the properties of the particular optical system. Nonetheless, an investigation of the advantages of the use of coordinates normally employed in Hamilton's method might prove most useful. For the resultant new feature would be the great simplicity of the identities between the aberration coefficients as defined above. The problem would then simply resolve itself into the calculation of the coefficients of the power series expansion of one or other of the characteristic functions of Hamilton, by means of an adaptation of these methods.

For references vide page 46 . (235)

PART II. THE GEOMETRICAL ABERRATIONS OF OPTICAL SYSTEMS WITH CONTINUOUSLY VARYING REFRACTIVE INDEX.

§10. Introduction.

(a) The theory of the aberrations of symmetrical optical systems consisting of a medium with continuously varying refractive index is of considerable heuristic interest. The aberrations may be obtained by considering such a system as the limiting case of a system of the type previously considered when the number of refracting surfaces tends to infinity, (v. §19(c)). A solution of very great simplicity may, however, be obtained more directly by dealing with the equations of variation arising from Fermat's principle; that is, by considering the optical analogue of Lagrange's Equations of analytical dynamics. The resulting development is in all respects analogous to that applying to a system of a finite number of refracting surfaces; a matter which will be more fully dealt with in §19. Here again it appears that for practical purposes at least these methods may be preferable to those of Hamilton. This is perhaps not altogether surprising. For in a great number of actual problems of analytical dynamics one naturally attempts a solution by means of Lagrange's equations (of the second kind), rather than by means of the canonical equations of Hamilton. Whereas the latter are frequently more suited to considerations of a more general or abstract kind, (e.g. the existence of algebraic integrals in the three body problem), they are not always convenient in the treatment of more 'practical' problems; and analogously in optics. Nevertheless it appears that so far the present subject has generally been dealt with by Hamiltonian methods, and then only as regards primary aberrations, (cf. Luneberg, 1944, Ch. V. p. 322). [Where 'Lagrangian' methods have been used, they have been, in the author's opinion, much encumbered by the use of cylindrical polar coordinates; cf. Picht, 1939, and the references given there.] Here, Cartesian coordinates are used throughout: and nothing prevents us from deriving higher order aberration coefficients in explicit form by an iterative process, (v. §17-18).

It may be added that apart from its heuristic value, the theory finds immediate application in connection with the electrostatic electron optical system (Frank und von Mises, 1943) (The electric-magnetic electron optical system may be dealt with by an extension of the method set out below. But on account of the anisotropy of the 'medium' the theory will be rather more complex.)

(b) It will be seen that the following method is really a special case of a much more general method of determining the solution of "nearly linear" second or higher order certain equations, or systems of such equations as series of ascending powers of initial values. The equations are called nearly linear if they can be regarded as perturbed linear equations, i.e. if the equations become linear when terms containing higher powers than the first of the dependent variables are neglected. (The perturbing terms must be capable of epansion as series in ascending powers of the dependent variables and their derivatives. The equations need not be derivable from a variational principle.

This method of solution is so 'obvious' that one would expect it to be a standard procedure; nonetheless the author has not been able to find any references to it in the literature on differential equations. (It may be noted that it is analogous to the method of variation of parameters for the determination of the particular integral of a non-homogeneous linear differential equation). However, it would lead too far to consider the general problem here.

§11. The 'continuous' symmetrical optical system.

(a) The passage of light rays through an optical system consisting of an isotropic medium the refractive index of which N is a function only (i) of x, (ii) of the perpendicular distance from the x-axis, (i.e. the axis of symmetry of the system), is governed by Fermat's Principle, in the form

$$\delta\int L dx = 0 \quad , \tag{11.1}$$

where the Lagrangian L is given by

$$L = N(\xi,x)\sqrt{1+\zeta} \quad , \tag{11.2}$$

with $\xi = Y^2 + Z^2$, $\eta = YV + ZW$, $\zeta = V^2 + W^2$, (11.21)

(x,Y,Z) constitutes an orthogonal Cartesian coordinate system, whilst $\underline{V} = \underline{Y}'$. Dashes now denote differentiation with respect to x, unless otherwise stated, and Clarendon type as usual indicates that the symbol represents both the Y and Z components of the quantity in question. (A table of symbols as used in this paper is appended in §20).

Write for the refractive index on the axis

$$N(0,x) = N_0(x) \quad . \tag{11.3}$$

Then we may define the "equi-refractive" ~~refractive~~ surfaces of the system by the equation

$$N(\xi,x+\bar{x}) = N_0(x) \quad , \tag{11.31}$$

where $\bar{X},Y,Z$ are current coordinates on such a surface, with $\underline{x}$ as origin, and over this surface N is constant. As usual we assume that the equi-refractive surfaces have no singularities in the vicinity of the $\underline{x}$-axis. Consequently we may write

$$N(\xi,x) = \sum_{n=0}^{\infty} N_n(x)\xi^n \quad , \tag{11.4}$$

the series on the right being convergent for relevant values of $\underline{x}$ and ξ. We now write

$$L = \sum_{n=0}^{\infty} L_n = L_0 + L_1 + L^* \, , \tag{11.5}$$

where L_n is a homogeneous polynomial of degree $\underline{n}$ in ξ,ζ , viz.

$$L_n = \sum_{m=0}^{\infty} \binom{\frac{1}{2}}{m} N_{n-m}(x)\xi^{n-m}\zeta^m \quad . \tag{11.51}$$

(The arguments of functions will usually be omitted when no confusion is likely thereby to arise.) In particular

$$\left.\begin{aligned}
L_0 &= N_0 \\
L_1 &= N_1\xi + \tfrac{1}{2}N_0\zeta \\
L_2 &= N_2\xi^2 + \tfrac{1}{2}N_1\xi\zeta - \tfrac{1}{8}N_0\zeta^2 \\
L_3 &= N_3\xi^3 + \tfrac{1}{2}N_2\xi^2\zeta - \tfrac{1}{8}N_1\xi\zeta^2 + \tfrac{1}{16}N_0\zeta^3 \, .
\end{aligned}\right\} \tag{11.52}$$

(b) Equation (11.1) gives the variational equations

$$[L]_{\underline{Y}} \equiv \frac{d}{dx}\frac{\partial L}{\partial \underline{V}} - \frac{\partial L}{\partial \underline{Y}} = 0 \quad . \tag{11.6}$$

If we write $\quad [L]_{\underline{Y}} = -\underline{F} \quad , \quad [L]_{\underline{Z}} = -G \quad ,$ (11.61)

this immediately leads to

$$(N_0\underline{V})' - 2N_1\underline{Y} = \underline{F} \quad . \tag{11.7}$$

§12. On the associated linear equation. p and q solutions.

(a) With (11.7) we associate the self-adjoint homogeneous linear differential equation

$$(N_0 v)' - 2N_1 y = 0 \quad , \tag{12.1}$$

where $\quad v = y' \quad .$ (12.11)

A definite solution of (12.1) requires the specification of two boundary conditions. We consider two particular linealy independent solutions of (12.1), corresponding respectively to the boundary conditions

$$\left.\begin{aligned} &\text{(i) at } x=a,\ y=1 \ , \quad v = 0 \\ &\text{(ii) at } x=a,\ y=0 \ , \quad v = 1 \quad . \end{aligned}\right\} \tag{12.2}$$

We distinguish the two solutions so obtained by the subscripts $\underline{p}$ and $\underline{q}$ respectively.

(b) Let $f(x)$ be one particular solution of (12.1). Then the general solution may be written as

$$\left.\begin{aligned} y &= Af(x) + Bf(x)\int_a^x \frac{dx}{[f(x)]^2 N_0(x)} \quad , \\ &\equiv f(x)\left[A + BJ(x)\right] \quad , \text{ say,} \end{aligned}\right\} \tag{12.3}$$

where $\underline{A}$ and $\underline{B}$ are constants of integration, and $\underline{a}$ is a constant. Making use of the boundary conditions (12.2) we find

$$y_p(x) = f(x)/f_1 - N_{01}f_1'f(x)J(x) \quad , \tag{12.4}$$

$$y_q(x) = N_{01}f_1 f(x)J(x) \quad , \tag{12.5}$$

$$v_p(x) = f'(x)/f_1 - N_{01}f_1'/N_0(x)f(x) - N_{01}f_1'f'(x)J(x) , \quad (12.6)$$

$$v_q(x) = N_{01}f_1/N_0(x)f(x) + N_{01}f_1f'(x)J(x) , \quad (12.7)$$

where the subscript '1' indicates the value of the function at $x = a$. The Wronskian (cf. Kamke, 1943, p.72),

$$\begin{vmatrix} y_p(x) & y_q(x) \\ v_p(x) & v_q(x) \end{vmatrix} = N_{01}/N_0(x) \neq 0 , \quad (12.8)$$

as is necessary if the p and q solutions are to be independent. Any particular solution of (12.1) may now be exhibited in the form

$$y(x) = y_p(x)y_1 + y_q(x)v_1 . \quad (12.9)$$

§13. On the solution curves of the linear equation.

(a) The equation (12.1) and its solution (12.9) completely govern the paraxial behaviour of the optical system, a particular solution curve of (12.9) representing a (tangential) paraxial ray. Thus consider a particular set of such curves all the members of which pass through the point (a,0); for these

$$y = y_q(x)v_1 . \quad (13.1)$$

Accordingly if we suppose that one of these curves ($v_1 \neq 0$) again intersects the axis at some point b , $b > a$, then all of them do, and b is determined by the condition $y_q(b) = 0$, or, by (12.5),

$$f(b)J(b) = 0 . \quad (13.2)$$

If we further exclude the uninteresting case that all possible solution curves pass through the point (b,0), then we must have $f(b) \neq 0$, since otherwise $y_p(b) = y_q(b) = 0$. Hence it follows that b is the solution of the algebraic equation

$$J(b) = 0 . \quad (13.3)$$

We assume hereafter that (13.3) has a root $> a$; and we shall generally consider only the first root, if there are several. The values of functions at $x=b$ will be distinguished by the subscript '2' .

For all curves passing through the point (a,y_1)

$$y(x) = y_p(x)y_1 + y_q(x)v_1 \quad , \tag{13.4}$$

where v_1 is a variable parameter. Hence all 'rays' pass through the point $(b, y_{p2}y_1)$. The (paraxial) magnification of the system associated with the two conjugate planes $x = a, b$ is therefore

$$m = y_{p2} = f_2/f_1 \quad . \tag{13.5}$$

(b) It is, of course, not essential that a itself should determine the position of the object plane. On the contary, let it be at $x = \bar{a}$, and let the conjugate image plane (in the sense above) be at $x = \bar{b}$. Then it is easily verified that $\bar{b}$ is given by the algebraic equation

$$\left.\begin{aligned} J(\bar{b}) &= J(\bar{a}) \quad , \\ \text{i.e.} \quad \int_{\bar{a}}^{\bar{b}} \frac{dx}{[f(x)]^2 N_o(x)} &= 0 \quad . \end{aligned}\right\} \tag{13.6}$$

For the paraxial magnification associated with these conjugate planes we then find in virtue of (13.6)

$$m = y_p(\bar{b})/y_p(\bar{a}) = f(\bar{b})/f(\bar{a}) \quad . \tag{13.7}$$

Notice that (12.8), viz.

$$N_o(x) \begin{vmatrix} y_p(x) & y_q(x) \\ v_p(x) & v_q(x) \end{vmatrix} = \text{const.} = N_{o1} \tag{13.8}$$

constitutes the extension of the Lagrange (Smith-Helmholtz) theorem, for the type of system here being considered.

§14. Introduction of $\underline{\Lambda}(x)$.

(a) After the preliminaries of §§ 11-13 we may now turn to the question of aberrations. Consider the expression

$$\underline{\Lambda}(x) = N_0(x)[y(x)\underline{V}(x) - v(x)\underline{Y}(x)] \quad . \qquad (14.1)$$

We have $\underline{\Lambda}' = (N_0\underline{V})'y + N_0\underline{V}v - (N_0v)'\underline{Y} - N_0v\underline{V}$.

In virtue of (11.7) and (12.1) it follows at once that

$$\underline{\Lambda}'(x) = y(x)\underline{F} \quad . \qquad (14.2)$$

Integration from $\underline{a}$ to $\underline{x}$ results in

$$\underline{\Lambda}(x) = \underline{\Lambda}_1 + \int_a^x y(t)\underline{F}(t)dt \quad , \qquad (14.3)$$

where $\underline{F}(x)$ are the functions which result if we substitute in $\underline{F}$ the solutions $\underline{Y}(x)$, $\underline{V}(x)$ of (11.7), (supposed known). We thus arrive at the pair of simultaneous equations

$$\left.\begin{aligned} \underline{\Lambda}_p &= N_0y_p\underline{V} - N_0v_p\underline{Y} = N_{01}\underline{V}_1 + \int_a^x y_p(t)\underline{F}(t)dt \\ \underline{\Lambda}_q &= N_0y_q\underline{V} - N_0v_q\underline{Y} = -N_{01}\underline{V}_1 + \int_a^x y_q(t)\underline{F}(t)dt \ . \end{aligned}\right\} \quad (14.4)$$

These may be solved for $\underline{Y}(x)$, $\underline{V}(x)$, making use of (2.6). We find

$$\left.\begin{aligned} \underline{Y}(x) &= y_p(x)[\underline{Y}_1 + \underline{\delta}_y(x)] + y_q(x)[\underline{V}_1 + \underline{\delta}_v(x)] \\ \underline{V}(x) &= v_p(x)[\underline{Y}_1 + \underline{\delta}_y(x)] + v_q(x)[\underline{V}_1 + \underline{\delta}_v(x)] \ , \end{aligned}\right\} \quad (14.5)$$

where

$$\left.\begin{aligned} \underline{\delta}_y(x) &= - N_{01}^{-1}\int_a^x y_q(t)\underline{F}(t)dt \\ \underline{\delta}_v(x) &= + N_{01}^{-1}\int_a^x y_p(t)\underline{F}(t)dt \ . \end{aligned}\right\} \quad (14.6)$$

The second member of (14.5) follows from the first of course also by simply differentiating the latter. Putting

$$\underline{F}(t) = -\frac{d}{dt}\frac{\partial L^*}{\partial \underline{V}} + \frac{\partial L^*}{\partial \underline{Y}} \quad , \qquad (14.7)$$

and integrating by parts, (14.6) becomes

$$\left.\begin{aligned}\delta_y(x) &= +\frac{1}{N_{01}}\left[y_q\frac{\partial L^*}{\partial \underline{V}}\right]_a^x - \frac{1}{N_{01}}\int_a^x\left(v_q\frac{\partial L^*}{\partial \underline{V}} + y_q\frac{\partial L^*}{\partial \underline{Y}}\right)dt\\ \underline{\delta}_v(x) &= -\frac{1}{N_{01}}\left[y_p\frac{\partial L^*}{\partial \underline{V}}\right]_a^x + \frac{1}{N_{01}}\int_a^x\left(v_p\frac{\partial L^*}{\partial \underline{V}} + y_p\frac{\partial L^*}{\partial \underline{Y}}\right)dt \quad .\end{aligned}\right\}(14.8)$$

§15. Aberrations in general.

We assume as before that the system produces a real image in the plane $x = b$ of an object in the plane $x = a$. Then, since $y_{q2} = 0$, the image height $\underline{Y}_2$ is given by

$$\underline{Y}_2 = y_{p2}(\underline{Y}_1 + \underline{\delta}_{y2}) \quad . \tag{15.1}$$

Hence, by (13.5), $\underline{Y}_2 = m(\underline{Y}_1 + \underline{\delta}_{y2})$. (15.2)

We have therefore for the aberration $\underline{\varepsilon}$ $(= \underline{Y}_2 - m\underline{Y}_1)$, by (14.6),

$$\underline{\varepsilon} = -\mu\int_a^b y_q(t)\underline{F}(t)dt \quad , \tag{15.3}$$

where $\mu = +m/N_{01}$. (15.31)

In virtue of (14.8) this may be rewritten in the form

$$\underline{\varepsilon} = -\mu\int_a^b\left(v_q\frac{\partial L^*}{\partial \underline{V}} + y_q\frac{\partial L^*}{\partial \underline{Y}}\right)dt \quad , \tag{15.4}$$

since the integrated part vanishes, $(y_{q1} = y_{q2} = 0)$. (15.4) may also be written as

$$\underline{\varepsilon} = -2\mu\int_a^b\left(y_q\underline{Y}\frac{\partial L^*}{\partial \xi} + v_q\underline{V}\frac{\partial L^*}{\partial \zeta}\right)dt \quad . \tag{15.5}$$

(15.4), or (15.5), together with (12.1) contains the whole story of the geometrical aberrations of the type of system under consideration, (for the conjugate planes $x = a,b$).

§16. Primary aberrations.

(a) The primary aberrations (for the conjugate planes $x = a,b$) follow directly from (15.5) if we take

$$L^* = L_2 = N_2\xi^2 + \tfrac{1}{2}N_1\xi\zeta - \tfrac{1}{8}N_0\zeta^2. \tag{16.1}$$

(Arbitrary planes of references may of course easily be dealt with). (15.5) gives

$$\underline{\varepsilon} = -\mu\int_a^b\left[y_q(4N_2\xi + N_1\zeta)\underline{Y} + v_q(N_1\xi - \tfrac{1}{2}N_0\zeta)\underline{V}\right]dt , \qquad (16.2)$$

where $\underline{Y}$, $\underline{V}$, ξ,ζ have their paraxial forms, i.e.

$$\underline{Y} = y_p(x)\underline{Y}_1 + y_q(x)\underline{V}_1, \quad \text{etc.} \qquad (16.21)$$

The expression (16.2) represents the complete monochromatic primary aberrations of the system, and may be used for practical purposes as it stands by carrying out the integrations necessary for the determination of the coefficients of $\underline{Y}_1\xi_1,\ldots,\underline{V}_1\zeta_1$.

If we write, for convenience,

$$\left.\begin{array}{l} P_1 = 4N_2y_p^4 , \quad P_2 = N_1y_p^2y_p^2 , \quad P_3 = -\tfrac{1}{2}N_0v_p^4 , \\ \text{and} \quad p = y_q/y_p , \quad q = v_q/v_p , \end{array}\right\} \qquad (16.22)$$

we obtain by inspection

$$\underline{\varepsilon} = \Gamma_1\underline{Y}_1\xi_1 + \Gamma_2\underline{V}_1\xi_1 + \Gamma_3\underline{Y}_1\eta_1 + \Gamma_4(2\underline{V}_1\eta_1 + \underline{Y}_1\zeta_1) + \Gamma_5\underline{V}_1\zeta_1, \qquad (16.3)$$

where

$$\left.\begin{array}{l} \Gamma_1 = -\mu\int_a^b[pP_1 + (p+q)P_2 + qP_3]dt \\ \Gamma_2 = -\mu\int_a^b[p^2P_1 + (p^2+q^2)P_2 + q^2P_3]dt \\ \Gamma_3 = -2\mu\int_a^b[p^2P_1 + 2pqP_2 + q^2P_3]dt \\ \Gamma_4 = -\mu\int_a^b[p^3P_1 + pq(p+q)P_2 + q^3P_3]dt \\ \Gamma_5 = -\mu\int_a^b[p^4P_1 + 2p^2q^2P_2 + q^4P_3]dt . \end{array}\right\} \qquad (16.4)$$

Considering $2\Gamma_2 - \Gamma_3$, we obtain

$$2\Gamma_2 - \Gamma_3 = -2\mu\int_a^b(p-q)^2P_2\,dt = -2\mu\int_a^b\frac{P_2N_{01}^2\,dt}{N_0^2y_p^2v_p^2} = -\mu N_{01}^2\varpi , \qquad (16.5)$$

by (12.8), where

$$\varpi = \int_a^b(2N_1/N_0^2)dt , \qquad (16.51)$$

so that we would normally calculate Γ_3 from Γ_2 by means of

$$\Gamma_3 = 2\Gamma_2 + \mu N_{01}^2\varpi. \qquad (16.52)$$

(b) In place of the equation (16.3) for ε one may use a form of $\underline{\varepsilon}$ very usual in connection with 'discontinuous' systems. It is derived from (16.3) by introducing 'entrance-pupil coordinates' τ, ψ, defined by

$$\left.\begin{aligned} V_1 &= \tau\cos\psi + sY_1 \\ W_1 &= \tau\sin\psi \quad , \end{aligned}\right\} \quad (16.6)$$

where $\underline{s}$ is a certain constant defining the position of the entrance pupil. (The object point lies in the tangential plane). We then immediately obtain the aberrations in the usual form

$$\left.\begin{aligned} \varepsilon_y &= \sigma_1\tau^3\cos\psi + \sigma_2\tau^2(2 + \cos2\psi) + (3\sigma_3 + \sigma_4)\tau\cos\psi + \sigma_5 \\ \varepsilon_z &= \sigma_1\tau^3\sin\psi + \sigma_2\tau^2\sin2\psi \qquad + (\sigma_3 + \sigma_4)\tau\sin\psi \quad , \end{aligned}\right\} (16.7)$$

with

$$\left.\begin{aligned} \sigma_1 &= \Gamma_5 \\ \sigma_2 &= (\Gamma_4 + s\Gamma_5)Y_1 \\ \sigma_3 &= (\tfrac{1}{2}\Gamma_3 + 2s\Gamma_4 + s^2\Gamma_5)Y_1^2 \\ \sigma_4 &= -\tfrac{1}{2}\mu N_{01}^2\varpi Y_1^2 \\ \sigma_5 &= [\Gamma_1 + s(\Gamma_2 + \Gamma_3) + 3s^2\Gamma_4 + s^3\Gamma_5]Y_1^3 \ . \end{aligned}\right\} \quad (16.8)$$

(c) (16.8) shows incidentally how the primary aberration coefficients behave when considered as functions of $\underline{s}$, that is of the position of the diaphragm;

$$\left.\begin{aligned} \partial\sigma_1/\partial s = \partial\sigma_4/\partial s &= 0 \\ \partial\sigma_2/\partial s &= \sigma_1 Y_1 \\ \partial\sigma_3/\partial s &= 2\sigma_2 Y_1 \\ \partial\sigma_5/\partial s &= (3\sigma_3 + \sigma_4)Y_1 \ . \end{aligned}\right\} \quad (16.9)$$

It will be seen that §§ 15 and 16 constitute the direct analogue of the usual Seidel theory of the primary aberrations.

§17. Aberrations of higher order. Iteration.

Equation (15.5) allows of the determination of the aberrations of any order by means of an iterative process. The latter corresponds very closely to that described in S§5. Thus, from (14.3) we have $\underline{\Lambda}(x) - \underline{\Lambda}_1 = -\left[y\frac{\partial L^*}{\partial \underline{V}}\right]_a^x + 2\int_a^x\left(y\underline{Y}\frac{\partial L^*}{\partial \xi} + v\underline{V}\frac{\partial L^*}{\partial \zeta}\right)dt$; (17.1) so that the 'p' and 'q' parts of this equation are essentially the increments (14.8). If on the right hand side of (17.1) we now substitute for L^* explicitly from (11.52), and then put $\underline{Y}(t) = y_p(t)\underline{Y}_1 + y_q(t)\underline{V}_1$ and $\underline{V}(t) = v_p(t)\underline{Y}_1 + v_q(t)\underline{V}_1$ we obtain what was previously called a 'partial expansion', but this time of the increments themselves. That is, (17.1) corresponds exactly to the sum, (v. S§5(a)),

$$\sum_j \Delta\underline{\Lambda}_j = \sum_j \sum_{n=1}^{\infty} \underline{\Gamma}_{n,j} \quad , \tag{17.2}$$

where $\underline{\Gamma}_{n,j}$ is given by S(5.21). Omitting from (17.1) all terms $O(5)$ we are left with the correct primary terms of the increments (When these have been obtained we have effectively the primary aberration coefficients for arbitrary positions of the object). Next, take (17.1) with $L^* = L_2 + L_3$. Make the substitutions

$$\left.\begin{aligned} \underline{Y}_1 &\rightarrow \underline{Y}_1 + \underline{\delta}_y(t) \\ \underline{V}_1 &\rightarrow \underline{V}_1 + \underline{\delta}_v(t) \quad , \end{aligned}\right\} \tag{17.3}$$

(the increments being correct as regards first order terms), and omit all terms $O(7)$. This leaves us with the correct primary and secondary aberration coefficients, for arbitrary object position. The way in which the tertiary, quaternary,... coefficients are to be obtained is now obvious. (Note that at each step we need of course only determine the terms of the corresponding order, all earlier terms remaining unchanged). If we proceed up to order $\underline{n}$, say, we may at the last step use equation (15.5), if the object position be fixed. We thereby avoid one series of integrations, viz. those giving the n'th order terms in (15.5) with y_q, v_q replaced by y_p, v_p.

In practical work one would naturally begin by constructing 'computing schemes' adequate to the problem in hand; and tables analogous to $\underline{S}$(21.2) may be prepared.

§18. Secondary aberrations.

In this section we briefly consider secondary aberrations, as a special case of the considerations of the preceding section. To determine them it is necessary first to find the primary increments from (14.8). The integrated part, of course, gives no trouble. Thus consider the expression

$$[y\underline{V}(N_1\xi - \tfrac{1}{2}N_0\zeta)]/N_{01}\ , \qquad (18.1)$$

and write $\underline{I}_q(x)$, $\underline{I}_p(x)$ for the integrated parts of (14.8). Then the latter are obtained by making the substitution (17.3) in (18.1) and integrating between the limits $\underline{a}$ and $\underline{x}$.

The integrals themselves are not difficult to deal with. In conformity with (16.4) we define the six integrals

$$\left.\begin{aligned} \Upsilon_0(x) &= N_{01}^{-1}\int_a^x [P_1 + 2P_2 + P_3]\,dt \\ \Upsilon_1(x) &= N_{01}^{-1}\int_a^x [pP_1 + (p+q)P_2 + qP_3]\,dt\ , \end{aligned}\right\} \qquad (18.2)$$

and $\Upsilon_2(x), \ldots, \Upsilon_5(x)$ similarly, by analogy from (16.4), i.e.

$$\Upsilon_i(x) = -\frac{1}{m}\int_a^x \left(\frac{\partial T_i(a,b)}{\partial b}\right)_{b=t} dt \quad , (i=1,\dots,5). \qquad (18.3)$$

Then

$$\left.\begin{aligned} \underline{\delta}_y(x) &= +\underline{I}_q(x) - [\Upsilon_1(x)\underline{Y}_1\xi_1 + \Upsilon_2(x)\underline{V}_1\xi_1 + \Upsilon_3(x)\underline{Y}_1\eta_1 + \\ &\qquad + \Upsilon_4(x)(2\underline{V}_1\eta_1 + \underline{Y}_1\zeta_1) + \Upsilon_5(x)\underline{V}_1\zeta_1] \ , \\ \underline{\delta}_v(x) &= -\underline{I}_p(x) + [\Upsilon_0(x)\underline{Y}_1\xi_1 + \Upsilon_1(x)(\underline{V}_1\xi_1 + 2\underline{Y}_1\eta_1) + \Upsilon_3(x)\underline{V}_1\eta_1 \\ &\qquad + \Upsilon_2(x)\underline{Y}_1\zeta_1 + \Upsilon_4(x)\underline{V}_1\zeta_1] \ . \end{aligned}\right\} \quad (18.4)$$

Now put $L^* = L_2 + L_3$ in (16.5). Then

$$\underline{\varepsilon} = -\mu\int_a^b \{ y_q\underline{Y}[4N_2\xi + N_1\zeta + 6N_3\xi^2 + 2N_2\xi\zeta - \tfrac{1}{4}N_1\zeta^2$$
$$+ v_q\underline{V}[N_1\xi - \tfrac{1}{2}N_0\zeta + N_2\xi^2 - \tfrac{1}{2}N_1\xi\zeta + \tfrac{3}{8}N_0\zeta^2]\}dt \ . \quad (18.5)$$

In the terms of the integrand quadratic in ξ,ζ we insert the paraxial solutions for $\underline{Y}(t)$, $\underline{V}(t)$; whilst in the terms linear in ξ,ζ we make the substitution (17.3), where the increments are given by (18.4). The resulting integral represents the complete primary and secondary aberrations of the system, for the conjugate planes x = a,b. Separate integrals may again be written down to obtain $\underline{\varepsilon}$ in the form similar to (16.3).

§ 19. On the limiting case of a k surface system, as $k \to \infty$.

(a)(i) As was mentioned in the introduction (§10(a)) we may regard a system with continuously varying refractive index as the limiting case of a system with k refracting surfaces, as $k \to \infty$, where at the same time $\lim_{k\to\infty}\Delta N_j = 0$. [At the risk of creating some confusion we have to use the notation of $\underline{S}$ and $\underline{A}$ when dealing with discontinuous systems, and the present notation for continuous systems. We use the symbol '$\sim$' to stand for the phrase 'corresponds to'. Thus,

$$\left.\begin{aligned} N &\sim N_0(x) \\ M_{j+1} - M_j &\sim dx\frac{dM(x)}{dx}\ , \text{ (for any 'M')} \\ \underset{\sim}{V} &\sim -\underset{\sim}{V} \\ u &\sim -v\ , \quad \text{etc.,} \end{aligned}\right\}\ (19.1)$$

quantities on the right hand side of such 'equations' always refer to the continuous system.]

(ii) We obtain the equation of an equi-refractive surface from (11.31) if we set

$$\bar{x} = \sum_{n=1}^{\infty} \theta_n \xi^n\ , \qquad (19.2)$$

in conformity with A(2.2). Thus, expanding $N(\xi, x+\bar{x})$ in a series of ascending powers of ξ and $\bar{x}$, we obtain in view of (11.4) for the first two θ's:

$$\left.\begin{aligned} \theta_1 &= -N_1/N'_0 \\ \theta_2 &= (N'_0)^{-1}[-N_2 + \tfrac{1}{2}(N_1^2/N'_0)'] \ . \end{aligned}\right\}\ (19.21)$$

The paraxial radius of curvature of the equi-refractive surface is accordingly

$$r_0 \sim -N'_0/2N_1\ . \qquad (19.22)$$

Now for the paraxial ray-tracing equations we have

$$\left.\begin{aligned} u_{j+1} &= (1-k_j)y_j/r_{0j} + k_j u_j \\ y_{j+1} &= y_j - d'_j u_{j+1}\ . \end{aligned}\right\}\ (19.23)$$

But

$$\left.\begin{aligned} (1-k_j) &\sim (N'_0/N_0)dx \\ d'_j &\sim dx \\ u_{j+1} - u_j &\sim -v'\,dx \\ y_{j+1} - y_j &\sim y'\,dx \end{aligned}\right\}\ (19.24)$$

Substituting these in (19.23) they become, respectively,

$$\left.\begin{aligned} (N_0 v)' - 2N_1 y &= 0 \\ y' - v &= 0\ , \end{aligned}\right\}\ (19.25)$$

(iii) In the case of the usual expression for Petzval curvature we find similarly

$$\sum_{j=1}^{k} r_{0j}\Delta(1/N_j) \sim \int_a^b (-N_0'/N_0^2)(-2N_1/N_0')dx = \int_a^b (2N_1/N_0^2)dx$$
$$= \varpi \quad , \text{ by } (16.51)\text{, } (19.26)$$

(b)(i) The examination of certain non-paraxial equations in this context is of some heuristic interest. The 'correspondence' of the various quantities involved is now somewhat more difficult to ascertain. The main difficulties arise from the fact that the two $\underline{Y}$'s do not correspond. Considering any point x,Y,Z on the ray, its slope there is given by $\underline{V}$, so that the two $\underline{V}$'s correspond, i.e. with the present sign convention

$$\underline{V} \sim -\underline{V} . \tag{19.31}$$

On the other hand the present $x,\underline{Y}$ clearly specify the coordinates of the point of incidence on the particular equi-refractive surface passing through this point. Its polar tangent plane passes through the axial point $x-\bar{x}$, where $\bar{x}$ is the solution of the equation $N_0(x-\bar{x}) = N(\xi,x)$. (19.32)

Hence with this value of $\bar{x}$,

$$\tilde{\underline{Y}} \sim \underline{Y} - \bar{x}\underline{V} \quad . \tag{19.33}$$

(ii) Since, by (14.2), $\underline{\Lambda}'(x)$ splits up into two factors, one of which does not contain paraxial quantities explicitly, such a splitting up might also be expected to occur if we consider

$$\Delta\tilde{\underline{\Lambda}}_j = \Delta N_j(y_j\underline{V}_j - u_j\tilde{\underline{Y}}_j) \quad , \tag{19.41}$$

where $\tilde{\underline{Y}}_j$ are the coordinates of the points of incidence of the ray on the j-th surface. In fact, we have at once, omitting the subscript j, $\quad \Delta\tilde{\underline{\Lambda}} = y\Delta N\underline{V} - \tilde{\underline{Y}}y\Delta N/r_0$, (19.42)

$$\left.\begin{array}{ll}\text{since} & \Delta y = \Delta\tilde{\underline{Y}} = 0 \\ \text{and} & \Delta Nu = (\Delta N)y/r_0 \quad .\end{array}\right. \tag{19.43}$$

Also, $\underline{\tilde{Y}} = \underline{Y} - x\underline{V}$, so that, by A(3.8), this may be put into the form $\quad \Delta\underline{\tilde{\Lambda}} = y(x\Delta N\underline{V} - \Delta\underline{C})/r_0$, (19.44)

which clearly exhibits its quasi-invariant nature. The $\underline{F}$ of (14.2) may therefore be regarded as corresponding to the coeffient of $\underline{y}$ in (19.44). [Unfortunately the simplicity of (19.44) is adversely overcompensated by the fact that

$$\tilde{\Lambda}_{j+1} - \tilde{\Lambda}'_j = N'_j u'_j V'_j (x_{j+1} - x_j) \neq 0 \text{ , in general.]} \qquad (19.45)$$

(c) Not only the paraxial relations but also the aberrations of the continuous system can be obtained by limiting processes. Some curious features then emerge. As an example we merely consider the primary spherical aberration coefficient σ_1. By (16.4) this is $\quad \sigma_1 = -\mu\int_a^b (N_2 y_q^4 + 2N_1 y_q^2 v_q^2 - \frac{1}{2}N_0 v_q^4)dx$. (19.51)

Since in the subsequent development all paraxial quantities appear with the subscript $\underline{q}$ we shall consistently suppress the latter. The spherical aberration of the discontinuous system, on the other hand, is given by (Buchdahl, 1948(1), equ.(7.2))

$$\sigma_1 = (1/N_k u_k)\sum_{j=1}^{k}[\tfrac{1}{2}N(1-k)yi^2(i'-u) + 4(N'-N)(\theta_2 - \theta_1^3)y^4]_j \quad .(19.52)$$

Inserting in (19.52) the various 'correspondences', (paying due attention to different sign conventions), it becomes

$$\sigma_1 = \mu\int_a^b\Big[\tfrac{1}{2}N'_0 y\left(\frac{N'_0 v - 2N_1 y}{N'_0}\right)^2\left(\frac{2N'_0 v - 2N_1 y}{N'_0}\right) + 4N'_0\, y^4\left(-\frac{N_2}{N'_0} + \frac{N_1 N'_1}{N_0'^2} - \frac{N_1^2 N''_0}{2N_0'^3} + \frac{N_1^3}{N_0'^3}\right)\Big]dx \quad . \qquad (19.54)$$

Multiplying out the various factors we find

$$\sigma_1 = \mu\int_a^b\Big\{-4N_2 y^4 - 5N_1 y^2 v^2 + N'_0 y v^3 + (8N_1^2/N'_0)y^3 v + [(4N_1 N'_1/N'_0) - (2N_1^2 N''_0/N'_0)]y^4\Big\}dx \quad . \qquad (19.6)$$

Now the fourth, fifth, and sixth terms of the integrand do not contribute to the integral, for they may be written as

$$(2N_1 y^4/N_0')' \quad . \tag{19.61}$$

The intgral of this vanishes at both limits, since $y_{q1}=y_{q2}=0$. Similarly, integrating the third term by parts, we find

$$\begin{aligned}\int_a^b N_0' y v^3 dx &= -\int_a^b (N_0 v^4 + 3N_0 y v^2 v')dx \\ &= -\int_a^b N_0 v^4 dx + 3\int_a^b y v^2 (N_0' v - 2N_1 y)dx \quad , \text{ by (12.1).}\end{aligned}$$

The integral on the left is the same as the second integral on the right. Hence we have

$$\int_a^b N_0' y v^3 dx = \int_a^b (\tfrac{1}{2} N_0 v^4 + 3N_1 y^2 v^2)dx \quad . \tag{19.62}$$

Substituting this in (19.6) we obtain

$$\sigma_1 = -\mu \int_a^b (4N_2 y^4 + 2N_1 y^2 v^2 - \tfrac{1}{2} N_0 v^4)dx \quad , \tag{19.7}$$

which is identical with (19.51).

The curious feature revealed by this development is that the integrand on the right hand side of (19.6) (which corresponds to the usual sum for the primary spherical aberration) is so complex, especially as regards the appearance of first and second derivatives of the $N_m(x)$. These are all absent in the final form (19.7). That this is so is explained by the possiblity of carrying out one general integration by parts in (15.3), viz. of the term $y \frac{d}{dt} \frac{\partial L^*}{\partial Y}$, together with the vanishing of the integrated part. (One would have thought that there might be a corresponding simplification in the expressions of the theory of discontinuous systems. But this does not appear to be the case).

§20. Conclusion.

The theory outlined above possesses such simplicity that it should be well suited to practical application, e.g. in electron optics. It is valuable from a didactic view-point as regards the light it sheds on theory of systems of refracting surfaces, as well as in respect of of its directness in the application of the differential equations corresponding to Lagrange's equations of motion.

It may be objected that little mention has been made of the convergence of the many power series which occur, or of the iteration itself. This, however, appears to be a general fault of optical theories. To the author's knowledge no formal proof has ever been given that even the simplest series converge, such as the series for the spherical aberration of a system of coaxial spherical refracting surfaces, that is, the series the coefficient of the first term of which is the first Seidel sum σ_1. It is quite possible that it may not be too difficult to establish the convergence of such series for continuous systems, by means of the usual devices of mean value theorems and dominant series, and the convergence of the corresponding series for discontinuous systems might then be deducible by using suitable limiting processes. This, however, must remain a subject for later investigation.

The following is a list of the more important symbols as used in this paper. The numbers in brackets again refer to the section in which the symbol first appears.

Symbol	Meaning of symbol
a	defines the position of the object plane. (12)
b	defines the position of the image plane. (12)
Γ_i	coefficients in first order terms of $\underline{\varepsilon}$. (16)
$\underline{\delta}_y(x)$, $\underline{\delta}_v(x)$	$-N_{01}^{-1}\int_a^b y_q\underline{F}dt$, $+N_{01}^{-1}\int_a^b y_p\underline{F}dt$. (14)
$\underline{\varepsilon}$	$\underline{Y}_2 - m\underline{Y}_1$. (15)
F	variational derivative of $-L^*$ w.r.to Y . (11)
η	$YV + ZW$. (11)
G	variational derivative of $-L^*$ w.r.to Z . (11)
ζ	$V^2 + W^2$. (11)
$\underline{I}(x)$	integrated parts of increments. (18)
$J(x)$	$\int_a^b [N_0(x)\{f(x)\}^2]^{-1}dx$. (12)
L	$N(\xi,x)\sqrt{1+\zeta}$. (11)
L_n	$\underline{n}$-th order term of L . (11)
L^*	$L - L_0 - L_1$. (11)
$\underline{\Lambda}(x)$	$N_0(x)[y(x)\underline{V}(x) - v(x)\underline{Y}(x)]$. (14)
μ	m/N_{01} . (16)
$N(\xi,x)$	refractive index. (11)
ξ	$Y^2 + Z^2$. (11)
$N_n(x)$	coefficient of ξ^n in expansion of $N(\xi,x)$. (11)
p (subscript)	refers to solution fixed by $x=a, y=1, v=0$. (12)
P_i	quantities occurring in the Γ integrals. (16)
p	y_q/y_p . (16)
q (subscripts)	refers to solution fixed by $x=a, y=0, v=1$. (12)
q	v_q/v_p . (16)

t	auxiliary variable of integration. (14)
$\underline{V}$	$\underline{Y}'$. (11)
v	y' . (12)
x,Y,Z	orthogonal Cartesian coordinates. (11)
y	paraxial Y coordinate. (12)

- - - - - - - - - - - - - - - -

'	denotes differentiation with respect to $\underline{x}$. (11)
1,2 (subscripts)	refer to object and image planes resp^{y}. (12,13)

For references see page 46. (235)

PART III. ON THE SINE RELATION.

§21. The case of systems of coaxial refracting surfaces of revolution.

The sine relation states that if, in the absence of spherical aberration, a symmetrical optical system forms a well-defined image h' of an infinitesimal object h situated on the axis of the system and perpendicularly to it, then for *any* ray passing through object and image

$$Nh\sin\theta = N'h'\sin\theta' \quad , \tag{21.1}$$

where θ, θ' are the inclinations of the ray to the axis in the object and image space respectively, the object being immersed in a medium of refractive index N and the image in a medium of refractive index N'. (cf. Steward, 1928, p.51.).

We propose in this section to give a very simple (algebraic) proof of (21.1) for the case of a system of coaxial refracting surfaces of revolution. As was shown in A§3, it follows immediately from elementary considerations that for such a system the quantity

$$K^* = N(\gamma H_y - \beta H_z) \tag{21.2}$$

is an optical invariant, so that

$$N(\gamma H_y - \beta H_z) = N'(\gamma' H'_y - \beta' H'_z) \quad , \tag{21.3}$$

where undashed and dashed quantities refer to object and image space respectively throughout. If the system produces a perfect image,

$$\underline{H}' = m\underline{H} \quad , \tag{21.4}$$

where m is the (paraxial) magnification. Hence (21.3) becomes

$$(N\beta - mN'\beta')/H_y = (N\gamma - mN'\gamma')/H_z \quad (= \text{const.} + O(2)) \ , \tag{21.5}$$

where the member of this equation in brackets easily follows by considering paraxial values of the ratios on the left. Hence, for infinitesimal $\underline{H}$,

$$N\underline{\beta} = mN'\underline{\beta}' \quad , \tag{21.6}$$

or, squaring and adding the component equations,

$$N\sqrt{1-\alpha^2} = mN'\sqrt{1-\alpha'^2} \quad . \tag{21.7}$$

But $\alpha = \cos\theta$, $\alpha' = \cos\theta'$, and $m = h'/h$. (21.8)

Consequently $Nh\sin\theta = N'h'\sin\theta'$,

which is (21.1).

§22. Generalisation of K^*.

By a method analogous to that of the preceding section the relation (21.1) may be established more generally for any (refracting) system whatever, provided it possesses axial symmetry. Not only may the refracting medium be heterogeneous and anisotropic, but it may be such that a ray 'instantaneously' in a plane containing the axis of symmetry of the system does not remain in that plane (as in the case of the electric-magnetic electron optical system). To prove the sine relation in this case we first derive in this section an expression for the generalisation of the optical invariant K^*.

The path of a ray of light through the type of system now under consideration is governed by Fermat's Principle in the form

$$\delta\int L(\xi,\eta,\zeta;x)dx = 0 \quad , \tag{22.1}$$

where $\xi = Y^2+Z^2$, $\eta = YV+ZW$, $\zeta = V^2+W^2$, and $\underline{V} = d\underline{Y}/dx$. (22.11) x,Y,Z constitute an orthogonal Cartesian coordinate system, the x-axis coinciding with the axis of symmetry of the optical system. (The Lagrangian L can always be written as a function of ξ,η,ζ,x. For, by hypothesis, L is invariant with reppect to the group of rotations which leave the x-axis invariant. Considering infinitesimal rotations the condition of invariance leads to a linear first order partial differential equation with four independent variables Y,Z,V,W, the general solution of which is an arbitrary function of three arguments which may be takeh to be ξ,η,ζ. The fifth variable x does not enter into the equation. Hence L may be taken to have the form $L(\xi,\eta,\zeta;x)$, where L need not be capable of expansion in ascending powers of ξ,η,ζ.)

The equations of the ray trajectories are

$$\frac{d}{dx}\frac{\partial L}{\partial \underline{V}} - \frac{\partial L}{\partial \underline{Y}} = 0 , \tag{22.2}$$

or

$$\frac{d}{dx}\left(\underline{Y}\frac{\partial L}{\partial \eta} + 2\underline{V}\frac{\partial L}{\partial \zeta}\right) - \left(2\underline{Y}\frac{\partial L}{\partial \xi} + \underline{V}\frac{\partial L}{\partial \eta}\right) = 0 . \tag{22.21}$$

This may be written in the form

$$\left.\begin{aligned} \frac{d}{dx}\left(2V\frac{\partial L}{\partial \zeta}\right) + Y\left(\frac{d}{dx}\frac{\partial L}{\partial \eta} - 2\frac{\partial L}{\partial \xi}\right) &= 0 \\ \frac{d}{dx}\left(2W\frac{\partial L}{\partial \zeta}\right) + Z\left(\frac{d}{dx}\frac{\partial L}{\partial \eta} - 2\frac{\partial L}{\partial \xi}\right) &= 0 . \end{aligned}\right\} \tag{22.22}$$

Multiplying the first of these by Z and the second by Y and subtracting

$$Y\frac{d}{dx}W\frac{\partial L}{\partial \zeta} - Z\frac{d}{dx}V\frac{\partial L}{\partial \zeta} = 0 .$$

But since $\underline{V} = d\underline{Y}/dx$ this may be written

$$\frac{d}{dx}(YW - ZV)\frac{\partial L}{\partial \zeta} = 0 ,$$

which on integration gives

$$K^* \equiv \Gamma(YW - ZV) = \text{const.}, \tag{22.3}$$

where $$\Gamma = \frac{\partial L}{\partial \zeta} \quad . \tag{22.31}$$

(In dynamics equation (22.3) plays the part of an integral of angular momentum). (22.3) constitutes the desired generalisation of the earlier K^* of equ.(21.2), and reduces to the latter if we set (cf. Buchdahl, 1948(4), equ.(11.2))

$$L = N(\xi,x)\sqrt{1+\zeta} \quad . \tag{22.32}$$

§23. The sine relation in the general case.

In view of the result (22.3) we may now derive the sine relation as in §21. If Y,Z and Y',Z' are the 'components of the object and image' respectively and m is again the paraxial magnification, then

$$\left.\begin{aligned} Y' &= m(Y\cos\psi - Z\sin\psi) \\ Z' &= m(Y\sin\psi + Y\cos\psi) \quad , \end{aligned}\right\} \tag{23.1}$$

where ψ is the angle between object and image (both being supposed 'linear'). Hence, by (22.3),

$$\Gamma(YW - ZV) = m\Gamma'[(Y\cos\psi - Z\sin\psi)W' - (Y\sin\psi + Z\cos\psi)V'] . \tag{23.2}$$

Therfore $$\frac{\Gamma V - m\Gamma'(V'\cos\psi + W'\sin\psi)}{Y} = \frac{\Gamma W - m\Gamma'(-V'\sin\psi + W'\cos\psi)}{Z} \tag{23.3}$$

When the object height tends to zero, $\underline{Y} \rightarrow 0$, $(Y^2+Z^2= h^2)$, so that, similarly to the case of (21.5)

$$\left.\begin{aligned} \Gamma V &= m\Gamma'(V'\cos\psi + W'\sin\psi) \\ \Gamma W &= m\Gamma'(-V'\sin\psi + W'\cos\psi) \quad . \end{aligned}\right\} \tag{23.4}$$

Squaring and adding we have now

$$\Gamma\tan\theta = m\Gamma'\tan\theta' \quad . \tag{23.5}$$

This may be given a more interesting form as follows. If $N(\xi,\eta,\zeta,x)$ is the refractive index of the 'medium', then

$$L = N\sqrt{1+\zeta} \quad . \tag{23.6}$$

Hence $$\Gamma = N/2\sqrt{1+\zeta} + \sqrt{1+\zeta}\,\frac{\partial N}{\partial\zeta} \quad . \tag{23.61}$$

Now $\underline{Y} = 0$; so that it is useful to write for the axial refractive index $$N(0,0,\zeta;x) = \mu(\theta,x) \quad . \tag{23.7}$$

Here $\zeta = \tan^2\theta$, and thus we have

$$2\Gamma\tan\theta = \mu\sin\theta + \frac{\partial\mu}{\partial\theta}\cos\theta \quad . \tag{23.8}$$

If we now write h'/h for $\underline{m}$, (23.5) finally becomes

$$\mu h\sin\theta + h\left(\frac{\partial\mu}{\partial\theta}\right)\cos\theta = \mu' h'\sin\theta' + h'\left(\frac{\partial\mu}{\partial\theta}\right)'\cos\theta' \quad , \tag{23.9}$$

which constitutes the desired generalised sine relation.

References.

Buchdahl, H.A.,	1946	Proc.Phys.Soc., Vol.58, p. 545.
"	1948 (1)	J.O.S.A., Vol.38. No.1., p.14.
"	1948 (2)	This thesis, page 32.
"	1948 (3)	This thesis, page 148.
"	1948 (4)	This thesis, page 209.
Frank, von Mises,	1943	Die Differentialgleichungen der Physik, Vol.II. Ch.I., p.36.
Kamke, E.,	1943	Differentialgleichungen.
Luneberg, R.K.,	1944	Mathematical Theory of Optics, Brown University Lecture Notes.
Picht, J.,	1939	Elektronenoptik. §§ 4 and 30.
Smith, T.,	1948	Proc. Phys. Soc., Vol.60, p. 303.
Steward, G.C.,	1928	The Symmetrical Optical System, Cambridge Tract., No.25.

236

Proof

On Eddington's higher order equations of the gravitational field.

By H. A. Buchdahl, B.Sc., A.R.C.S.,

Physics Department, University of Tasmania.

(Communicated by G. C. McVittie, O.B.E., Ph.D.)

Received 14th January, 1947. *Read 17th January, 1947.*

Einstein's fundamental equations of the gravitational field are

$$G^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}G + \lambda g^{\mu\nu} = -\kappa T^{\mu\nu}, \ (\mu = 1, \ldots, 4) \tag{1}$$

where $T^{\mu\nu}$ are the components of the energy tensor and λ is the cosmical constant. In empty space these equations become

$$G^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}G + \lambda g^{\mu\nu} = 0, \tag{2}$$

which may be reduced to

$$G^{\mu\nu} = \lambda g^{\mu\nu} \tag{3}$$

since $G = 4\lambda$, by contraction of (2).
Eddington[1] has shown (§60, p. 138) that when the cosmological term in (2) is neglected these equations may be derived by Hamiltonian differentiation with respect to $g_{\mu\nu}$ of G, viz.

$$\frac{\partial G}{\partial g^{\mu\nu}} = -(G^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}G) = 0 \tag{4}$$

Distinguishing Eddington's equations and paragraphs by the letter E we have, by (E 35.3)

$$\delta\sqrt{-g} = \tfrac{1}{2}\sqrt{-g}\, g^{\mu\nu}\delta g_{\mu\nu}, \tag{5}$$

so that (2) may be obtained by variation of the integral

$$\int (G - 2\lambda)\sqrt{-g}\,d\tau, \tag{6}$$

i.e. by replacing G in (4) by

$$K = G - 2\lambda. \tag{6.1}$$

The spherically symmetrical solution of (3) may be written in the form (E 45.3)

$$ds^2 = -\gamma^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta\, d\phi^2 + \gamma dt^2 \\ \text{with } \gamma = 1 - 2m/r - \tfrac{1}{3}\lambda r^2 \tag{7}$$

where m is a constant of integration. When $\lambda = 0$ (7) reduces to the Schwarzschild solution for which

$$\gamma = 1 - 2m/r \tag{7.1}$$

Now Eddington has suggested (E § 6,2. p. 141) that instead of (4) the equations

$$\frac{\partial K}{\partial g_{\mu\nu}}=0 \tag{8}$$

might have been used, K being some fundamental invariant **other than** G. In particular he considers

$$K' = G_{\mu\nu}G^{\mu\nu} = g^{\alpha\beta}g^{\sigma\rho}g^{\mu\epsilon}g^{\nu\tau}B_{\mu\nu\sigma\rho}B_{\epsilon\tau\alpha\beta}, \tag{9.1}$$

$$K'' = B_{\mu\nu\sigma\rho}B^{\mu\nu\sigma\rho} = g^{\mu\alpha}g^{\nu\beta}g^{\sigma\epsilon}g^{\rho\tau}B_{\mu\nu\sigma\rho}B_{\alpha\beta\epsilon\tau}, \tag{9.2}$$

as being the simplest alternative invariants; but on these grounds we **should** also consider the square of the scalar **curvature**.

$$K''' = G^2 = g^{\mu\nu}g^{\sigma\rho}g^{\alpha\beta}g^{\epsilon\tau}B_{\mu\nu\sigma\rho}B_{\alpha\beta\epsilon\tau}. \tag{9.3}$$

Eddington shows that the Schwarzschild solution (7.1) is also a **solution of the alternative equations**

$$\frac{\partial K'}{\partial g_{\mu\nu}} = 0, \tag{10.1}$$

$$\frac{\partial K''}{\partial g_{\mu\nu}} = 0; \tag{10.2}$$

and it is easily seen **also** to be a solution of

$$\frac{\partial K'''}{\partial g_{\mu\nu}} = 0. \tag{10.3}$$

The author now considers these and other alternative field laws **and shows, amongst** other results, **that** in 4-space *every* solution **of (3)** is **also a solution of** (10.1) and (10.3), **and that** those **solutions of** (3) **which represent spaces of** constant Riemannian curvature **are also solutions of (10.2).** In particular it is verified that (7) (with λ **replaced by a constant of integration** a) **is not only a** solution of (10.1) and (10.3) **but also of (10.2) although** the space is not of **constant** Riemannian **curvature unless** $m = 0$.

(a) Consider a small variation of K'.

$$K' = \delta(G_{\mu\nu}G^{\mu\nu}) = G_{\mu\nu}\delta G^{\mu\nu} \quad G^{\mu\nu}\delta G_{\mu\nu}.$$

If

$$G_{\mu\nu} = a g_{\mu\nu}, \tag{11}$$

where a **is a constant,** then

$$\begin{aligned}
\delta K' &= a(g_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta G_{\mu\nu}) \\
&= a(\delta G - G^{\mu\nu}\delta g_{\mu\nu} + \delta G - G_{\mu\nu}\delta g^{\mu\nu}) \\
&= 2a\delta G - a^2(g^{\mu\nu}\delta g_{\mu\nu} + g_{\mu\nu}\delta g^{\mu\nu}) \\
&= 2a\delta G.
\end{aligned}$$

$$\begin{aligned}\therefore\ \delta\mathfrak{K}' = \delta(\sqrt{-g}K') &= \sqrt{-g}\,\delta K' + K'\delta\sqrt{-g}\\ &= 2a\sqrt{-g}\,\delta G + 4a^2\delta\sqrt{-g}\\ &= 2a(\delta\mathfrak{G} - G\delta\sqrt{-g}) + 4a^2\delta\sqrt{-g}\\ &= 2\sigma\delta(\mathfrak{G} - 2a\sqrt{-g}),\end{aligned}$$

i.e.
$$\delta\mathfrak{K}' = 2a\delta\mathfrak{K}. \tag{11.1}$$

Hence the condition $\int \delta\mathfrak{K}'d\tau = 0$ leads to the condition

$$\int \delta\mathfrak{K}d\tau = 0, \tag{11.2}$$

which is satisfied in virtue of the fact that equation (11) is equivalent to equation (11.2). Consequently
every solution of (11) *is a solution of the alternative equations* (10.1), where a is to be regarded as a constant of integration.

It is of interest to note that when $a \neq 0$ this result does not hold when the dimensional number is other than four. Thus in a space of $n(\neq 4)$ dimensions we have as before $\delta K' = 2a\delta G$;

but now
$$\begin{aligned}\delta\mathfrak{K}' &= \sqrt{-g}\,\delta K' + K'\delta\sqrt{-g}\\ &= 2a\sqrt{-g}\,\delta G + na^2\delta\sqrt{-g}\\ &= 2a(\delta\mathfrak{G} - G\delta\sqrt{-g}) + na^2\delta\sqrt{-g}\\ &= 2a\delta\mathfrak{K} + (n-4)a^2\delta\sqrt{-g},\end{aligned}$$

where the K of (8.1) is now replaced by K = G — (n — 2)c

Hence, if $\int \delta\mathfrak{K}\,d\tau = 0$ it follows that,

$$\int \delta\mathfrak{K}'d\tau = +(n-4)\,a^2\int\delta\sqrt{-g}\,d\tau, \tag{11.3}$$

which will in general fail to vanish.

As a particular case of the general result above we shall verify that (7), with
$$\gamma = 1 - 2m/r - a\cancel{r^2}, \tag{12}$$
(both m and a being constants of integration) is a solution of (10.1). To conform with Eddington's notation we write

$$e^{-\lambda} = e^{\nu} = \gamma = 1 - 2m/r - a\cancel{r^2} \tag{12.1}$$

where now, of course, α does not denote the cosmical constant. Using the form of the $G_{\mu\nu}$ given by (E 38.61-5) we find

$$\begin{aligned}\mathfrak{K}' = e^{-\frac{1}{2}\lambda+\frac{1}{2}\nu}\{&2r^2\Omega^2 + 2r(\nu'-\lambda')\Omega + \tfrac{1}{2}(3\nu'^2 - 2\nu'\lambda' + 3\lambda'^2)\\ &+ 2(e^{\lambda}-1)(\lambda'-\nu')/r + 2(e^{\lambda}-1)^2/r^2\}.\end{aligned} \tag{13}$$

Here $\Omega = \frac{1}{2}\nu'' - \frac{1}{4}\lambda'\nu' + \frac{1}{4}\nu'^2 (= g^{44}B_{1414})$, a dash denotes differentiation with respect to r, and we have set $\theta = \pi/2$ since it is not necessary to consider variations from the symmetrical condition

Using (12.1) after the various partial differentiations have been performed, we find, following some straightforward but tedious algebraic work, that

$$\left.\begin{aligned}
\frac{\partial K'}{\partial \lambda} &= -\tfrac{3}{2}K' + 2e^{-\frac{1}{2}\lambda+\frac{1}{2}\nu}\left\{(\lambda'-\nu')/r + 2(e^{\lambda}-1)/r^2\right\} = 4a - 6a^2r^2,\\
\frac{\partial K'}{\partial \lambda'} &= e^{-\frac{3}{2}\lambda+\frac{1}{2}\nu}\Big\{-\Omega(r^2\nu'+2r) + \tfrac{1}{2}r\nu'(\lambda'-\nu') + (3\lambda'-\nu')\\
&\qquad\qquad + 2(e^{\lambda}-1)/r\Big\} = 4ar - 6am - a^2r^3,\\
\frac{\partial K'}{\partial \lambda''} &= 0,\\
\frac{\partial K'}{\partial \nu} &= \tfrac{1}{2}K' = 2a^2r^2,\\
\frac{\partial K'}{\partial \nu'} &= e^{-\frac{3}{2}\lambda+\frac{1}{2}\nu}\Big\{\Omega[r^2(2\nu'-\lambda')+2r] + r(\nu'-\lambda')(\nu'-\tfrac{1}{2}\lambda')\\
&\qquad + (3\nu'-\lambda') - 2(e^{\lambda}-1)/r\Big\} = -4ar + 2am + \tfrac{10}{3}a^2r^3,\\
\frac{\partial K'}{\partial \nu''} &= e^{-\frac{3}{2}\lambda+\frac{1}{2}\nu}\left\{2r^2\Omega + r(\nu'-\lambda')\right\} = -2ar^2 + 4amr + \tfrac{2}{3}a^2r^4.
\end{aligned}\right\} \tag{13.1}$$

If these values are substituted in the expressions for the Lagrange derivatives

$$\left.\begin{aligned}
[K']_\lambda &= \frac{\partial K'}{\partial \lambda} - \frac{d}{dr}\frac{\partial K'}{\partial \lambda'} + \frac{d^2}{dr^2}\frac{\partial K'}{\partial \lambda''},\\
[K']_\nu &= \frac{\partial K'}{\partial \nu} - \frac{d}{dr}\frac{\partial K'}{\partial \nu'} + \frac{d^2}{dr^2}\frac{\partial K'}{\partial \nu''},
\end{aligned}\right\} \tag{13.2}$$

they will be found to vanish identically.

(b) Unfortunately it does not appear possible to deal with K'' in the manner above by using equation (11). But a somewhat more specialised result may be obtained by considering spaces of constant Riemannian curvature[2]. We therefore replace (11) by the equation

$$B_{\mu\nu\sigma\rho} = \tfrac{1}{3}a\,(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\sigma}g_{\nu\rho}) \tag{14}$$

Contracting for σ and ρ (11) is seen to be satisfied. Then

$$\begin{aligned}
\delta K'' &= B_{\mu\nu\sigma\rho}\delta B^{\mu\nu\sigma\rho} + B^{\mu\nu\sigma\rho}\delta B_{\mu\nu\sigma\rho}\\
&= \frac{a}{3}\left\{(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\sigma}g_{\nu\rho})\,\delta B^{\mu\nu\sigma\rho} + (g^{\mu\nu}g^{\sigma\rho} - g^{\mu\sigma}g^{\nu\rho})\,\delta B_{\mu\nu\sigma\rho}\right\}\\
&= \frac{a}{3}\left\{4\delta G - B^{\mu\nu\sigma\rho}\delta(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\sigma}g_{\nu\rho}) - B_{\mu\nu\sigma\rho}\delta(g^{\mu\nu}g^{\sigma\rho} - g^{\mu\sigma}g^{\nu\rho})\right\},
\end{aligned}$$

since $(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\sigma}g_{\nu\rho})\,B^{\mu\nu\sigma\rho} = g_{\mu\nu}g_{\sigma\rho}B^{\mu\nu\sigma\rho} + g_{\mu\sigma}g_{\nu\rho}B^{\mu\sigma\nu\rho} = 2G$,

$B^{\mu\nu\sigma\rho}$ being anti-symmetrical in ν and σ. As it is also anti-symmetrical in μ and ρ we get

$$\delta K'' = \frac{a}{3}(4\delta G - 4G^{\mu\nu}\delta g_{\mu\nu} - 4G_{\mu\nu}\delta G^{\mu\nu})$$

$$= \tfrac{4}{3}\, a\delta G - \tfrac{4}{3}a^2 (g^{\mu\nu}\delta g_{\mu\nu} + g_{\mu\nu}\delta g^{\mu\nu}) = \tfrac{4}{3}a\delta G.$$

Hence $\delta\mathfrak{K}'' = \sqrt{-g}\,\delta K'' + K''\delta\sqrt{-g} = \tfrac{4}{3}a\,(\sqrt{-g}\,\delta G + 2a\delta\sqrt{-g})$

$$= \tfrac{4}{3}a\,(\delta\mathfrak{G} - 2a\delta\sqrt{-g});$$

$$\therefore\ \delta\mathfrak{K}'' = \tfrac{4}{3}a\delta\mathfrak{K}. \qquad \textbf{(15)}$$

Therefore, as in the case of (11.1) we conclude that *any solution of* (11) *defining the line-element of a space of constant Riemannian curvature is also a solution of the alternative equations* (10.2).

Now (7) reduces to the de Sitter line-element (E 70.1) when $m = 0$, and the latter defines a space of constant Riemannian curvature. Hence we know now that (10.2) is satisfied by the two particular solutions

$$\gamma = 1 - 2m/r, \text{ and } \gamma = 1 - \tfrac{1}{3}ar^2.$$

It is natural to enquire therefore whether (12) as it stands may not perhaps also be a solution of (10.2). By direct substitution in the equations following (E 62.6) this may actually be verified to be the case. Thus using (12.1) again we find

$$\left.\begin{aligned}
\frac{\partial\mathfrak{K}''}{\partial\lambda} &= \frac{16m}{r^3} + \frac{8a}{3} - \frac{72m^2}{r^4} - 4a^2r^2,\\
\frac{\partial\mathfrak{K}''}{\partial\lambda'} &= -\frac{8m}{r^2} + \frac{8ar}{3} + \frac{24m^2}{r^3} - 4am - \frac{4a^2r^3}{3},\\
\frac{\partial\mathfrak{K}''}{\partial\lambda''} &= 0,\\
\frac{\partial\mathfrak{K}''}{\partial\nu} &= \frac{24m^2}{r^4} + \frac{4a^2r^2}{3},\\
\frac{\partial\mathfrak{K}''}{\partial\nu'} &= \frac{8m}{r^2} - \frac{8ar}{3} - \frac{40m^2}{r^3} + \frac{20am}{3} + \frac{20a^2r^3}{9},\\
\frac{\partial\mathfrak{K}''}{\partial\nu''} &= -\frac{8m}{r} - \frac{4ar^2}{3} + \frac{16m^2}{r^2} + \frac{16amr}{3} + \frac{4a^2r^4}{9}.
\end{aligned}\right\} \qquad \textbf{(16)}$$

With these values we find that the Lagrange derivatives $[\mathfrak{K}'']_\lambda$ and $[\mathfrak{K}'']_\nu$ vanish identically. Hence

(12) *is also a solution of the alternative equations* (10.2).

Comparing (13.1) and (16) we notice that when $m = 0$

$$\frac{[\mathfrak{K}'']_\lambda}{[\mathfrak{K}']_\lambda} = \frac{[\mathfrak{K}'']_\nu}{[\mathfrak{K}']_\nu} = \tfrac{2}{3},$$

as is required by (11.1) and (15), viz.

$$\frac{\delta\mathfrak{K}''}{\delta\mathfrak{K}'} = \tfrac{2}{3} \tag{16.1}$$

As in the case of K' the result established above (for $a \neq 0$) again requires the space to be four dimensional. The equation analogous to (11.3) is

$$\int \delta\mathfrak{K}'' d\tau = + \frac{2a^2(n-4)}{(n - .)} \int \delta\sqrt{-g}\, d\tau \tag{17}$$

(c) It is scarcely necessary to consider the case (10.3) in detail. It will be sufficient to quote the result that, assuming $G = na$,

$$\int \delta\mathfrak{K}''' d\tau = + na^2(n-4) \int \delta\sqrt{-g}\, d\tau \tag{17.1}$$

Hence *when* $n = 4$ *the result stated for* K' *apply here also.*

It may be remarked that $\mathfrak{K}'''$ is identical with Weyl's action density (E § 90. p. 209) in the absence of electromagnetic fields.

(d) The three alternative fundamental invariants which have been considered so far have in common the property of being quadratic in the second derivatives of the $g_{\mu\nu}$. It is of interest to examine whether the results established would also follow for more complex invariants. Consider for example the m^{th} power of the scalar curvature in 4-dimensional space. Let

$$K^{(m)} = G^m. \tag{18}$$

$$\text{Then } \delta K^{(m)} = mG^{m-1}\delta G = m(4a)^{m-1}\delta G$$

$$\begin{aligned}
\therefore \delta\mathfrak{K}^{(m)} &= \sqrt{-g}\,\delta K^{(m)} + K^{(m)}\delta\sqrt{-g} \\
&= m(4a)^{m-1}(\delta\mathfrak{G} - G\delta\sqrt{-g}) + (4a)^m\delta\sqrt{-g} \\
&= m(4a)^{m-1}\delta\left(\mathfrak{G} - \frac{4(m-1)}{m} a\sqrt{-g}\right).
\end{aligned}$$

Hence $\int \delta\mathfrak{K}^{(m)} d\tau = 0$ is consistent with (11.2) only if $m = 2$. i.e. in the case dealt with above. In particular we conclude that (12), ($a \neq 0$), is not a solution of $[K^{(m)}]_\lambda = [K^{(m)}]_\nu = 0$, except in that case. The author has considered various other invariants involving 3rd, 4th,... powers of the second derivatives of the $g_{\mu\nu}$, and these yielded similar results.

REFERENCES.

1. Eddington, 1930. *The Mathematical Theory of Relativity*, 2nd edn. C.U.P.
2. Eisenhart, 1926. *Riemannian Geometry*, § 26, p. 83. Princeton U.P.

THE HAMILTONIAN DERIVATIVES OF A CLASS OF FUNDAMENTAL INVARIANTS

By H. A. BUCHDAHL (*Hobart*)

[Received 27 August 1947]

1. THE Hamiltonian derivatives of fundamental invariants play a prominent part in the Theory of Relativity* since they are symmetrical fundamental tensors the divergence of which vanishes identically. In the present paper I consider the class of fundamental invariants which contain the derivatives of the $g_{\mu\nu}$ only up to the second order. In particular I shall establish amongst other results the following *explicit* formula for the Hamiltonian derivative $hK/hg_{\mu\nu}$ of such an invariant K

$$\frac{hK}{hg_{\mu\nu}} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3}B^{\mu}{}_{.\alpha\sigma\rho}Z^{\nu\alpha\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}K, \tag{1.1}$$

where $Z^{\mu\nu\sigma\rho}$ is the tensor $\partial K/\partial g_{\mu\nu,\sigma\rho}$ symmetrical in μ, ν and in σ, ρ, as may generally be determined by inspection when the form of K is known explicitly. Subscripts after a comma or a semicolon denote ordinary and covariant differentiation respectively, whilst the order of the subscripts of the covariant curvature tensor $B_{\mu\nu\sigma\rho}$ is indicated in (4.3) below. We can then verify directly that the divergence of the right-hand side of (1.1) vanishes identically; and the derivation of the Hamiltonian derivatives of various specific invariants is a matter of ease.

2. Since, by hypothesis, K contains the derivatives of the $g_{\mu\nu}$ only up to the second order, it follows from a well-known result of the calculus of variations that

$$P^{\mu\nu} \equiv \frac{hK}{hg_{\mu\nu}} = \frac{1}{\gamma}\left\{\frac{\partial(\gamma K)}{\partial g_{\mu\nu}} - \frac{\partial}{\partial x_\alpha}\frac{\partial(\gamma K)}{\partial g_{\mu\nu,\alpha}} + \frac{\partial^2}{\partial x_\alpha \partial x_\beta}\frac{\partial(\gamma K)}{\partial g_{\mu\nu,\alpha\beta}}\right\}, \tag{2.1}$$

where γ is written for $\sqrt{(-g)}$.

Now, if $\theta(K)$ is an arbitrary function of K (satisfying such conditions of continuity and differentiability as are implied in the following argument), then $\theta(K)$ is itself a fundamental invariant; and its Hamiltonian derivative is given by (2.1) if K be replaced by $\theta(K)$.

Write

$$\theta' = \frac{d\theta(K)}{dK}, \qquad \theta'' = \frac{d^2\theta(K)}{dK^2},$$

* Eddington (**1**), § 61, ff.

and so on. Then from (2.1) we have

$$\frac{h\{\theta(K)\}}{hg_{\mu\nu}} = \tfrac{1}{2}g^{\mu\nu}\theta + \theta'\left[\frac{\partial K}{\partial g_{\mu\nu}} - \left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}\right)_{,\alpha} + \frac{\gamma_{,\alpha}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha}} + \right.$$

$$\left. + \left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\alpha\beta} + \frac{2\gamma_{,\alpha}}{\gamma}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\beta} + \frac{\gamma_{,\alpha\beta}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right] +$$

$$+\theta''\left[-K_{,\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha}} + K_{,\alpha\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}} + 2K_{,\alpha}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right) + \frac{2\gamma_{,\beta}}{\gamma}K_{,\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right] +$$

$$+\theta''' K_{,\alpha}K_{,\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}. \quad (2.2)$$

$P^{\mu\nu}$ is then given by (2.2) if we choose $\theta(K) = K$; i.e.

$$P^{\mu\nu} = \tfrac{1}{2}g^{\mu\nu}K + \frac{\partial K}{\partial g_{\mu\nu}} - \left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}\right)_{,\alpha} - \frac{\gamma_{,\alpha}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha}} +$$

$$+\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\alpha\beta} + \frac{2\gamma_{,\alpha}}{\gamma} + \left(\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}\right)_{,\beta} + \frac{\gamma_{,\alpha\beta}}{\gamma}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}. \quad (2.3)$$

Now the right-hand side of (2.2) is a tensor. Since $\theta(K)$ is an arbitrary function of its argument,

$$K_{,\alpha}K_{,\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}} = K_{;\alpha}K_{;\beta}\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}$$

must be a tensor. Hence $\dfrac{\partial K}{\partial g_{\mu\nu,\alpha\beta}}$, which is symmetrical in α, β (and in μ, ν) is a tensor. We write

$$\frac{\partial K}{\partial g_{\mu\nu,\alpha\beta}} = Z^{\mu\nu\alpha\beta}, \quad (2.4)$$

where

$$Z^{\mu\nu\alpha\beta} = Z^{\mu\nu\beta\alpha} = Z^{\nu\mu\alpha\beta}. \quad (2.41)$$

In the same way the coefficient of θ'' in (2.3), i.e.

$$-K_{,\alpha}\frac{\partial K}{\partial g_{\mu\nu,\alpha}} + K_{,\alpha\beta}Z^{\mu\nu\alpha\beta} + 2K_{,\alpha}Z^{\mu\nu\alpha\beta}{}_{,\beta} + \frac{2\gamma_{,\beta}}{\gamma}K_{,\alpha}Z^{\mu\nu\alpha\beta} \quad (2.5)$$

is a tensor. But

$$K_{,\alpha\beta} = (K_{;\alpha})_{,\beta} = K_{;\alpha\beta} + \{\alpha\beta,\sigma\}K_{;\sigma}, \quad (2.6)$$

$$\gamma_{,\beta}/\gamma = \{\beta\sigma,\sigma\}, \quad (2.61)$$

and

$$Z^{\mu\nu\alpha\beta}{}_{,\beta}$$

$$= Z^{\mu\nu\alpha\beta}{}_{;\beta} - \{\sigma\beta,\mu\}Z^{\sigma\nu\alpha\beta} - \{\sigma\beta,\nu\}Z^{\mu\sigma\alpha\beta} - \{\sigma\beta,\alpha\}Z^{\mu\nu\sigma\beta} - \{\sigma\beta,\beta\}Z^{\mu\nu\alpha\sigma}. \quad (2.62)$$

Substituting in (2.5), using (2.41), and changing dummy indices we get

$$K_{;\alpha\beta}Z^{\mu\nu\alpha\beta}+2K_{;\alpha}Z^{\mu\nu\alpha\beta}{}_{;\beta}-$$
$$-K_{;\alpha}\left(\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}+2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}+\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho}\right). \quad (2.7)$$

Hence the quantity

$$Y^{\mu\nu\alpha}=\frac{\partial K}{\partial g_{\mu\nu,\alpha}}+2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}+2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}+\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho} \quad (2.8)$$

is a tensor.

3. Now without loss of generality we may for the time being (§§ 3–7) restrict ourselves to the case in which K has the form

$$K=\Gamma^{\mu_1\nu_1\ldots\rho_m}\prod_{s=1}^{m}B_{\mu_s\nu_s\sigma_s\rho_s}, \quad (3.1)$$

where $\Gamma^{\mu_1\nu_1\ldots\rho_m}$ is a tensor of rank $4m$ made up entirely of the product of $2m$ components of the metrical tensor. For the most general second-order invariant is of the form

$$K=F(K_1,K_2,\ldots), \quad (3.11)$$

where F is some function of the K_1, each of the latter being of the form (3.1). When the Hamiltonian derivatives of these are known, that of K follows at once, and is given by (1.1), (as is seen in § 8).

Fundamental *relative* invariants may also be dealt with by the present method. These are formed in a V_n with the help of the numerical tensor densities $\epsilon^{\mu_1\mu_2\ldots\mu_n}$ and $\epsilon_{\mu_1\mu_2\ldots\mu_n}$ [(**3**), 25]. How to determine their Hamiltonian derivatives is best illustrated by a simple example. Let $n=2$ and suppose that the Hamiltonian derivative of

$$\mathfrak{K}=\sqrt{(-\epsilon^{\mu\sigma}\epsilon^{\nu\rho}G_{\mu\nu}G_{\sigma\rho})} \quad (3.2)$$

be required. Raising the subscripts ν and ρ, we get

$$\mathfrak{K}=\sqrt{(-\epsilon^{\mu\sigma}\epsilon^{\nu\rho}g_{\nu\alpha}g_{\rho\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})}. \quad (3.3)$$

But $$\epsilon^{\nu\rho}g_{\nu\alpha}g_{\rho\beta}=\epsilon_{\alpha\beta}g, \quad [(\mathbf{3})\ 5,\ (3.43)]. \quad (3.31)$$

Hence $$\mathfrak{K}=\gamma\sqrt{(\epsilon^{\mu\sigma}\epsilon_{\alpha\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})}=\gamma\sqrt{(\delta^{\mu\sigma}_{\alpha\beta}G^{\alpha}_{\mu}G^{\beta}_{\sigma})},$$

where $\delta^{\mu\sigma}_{\alpha\beta}$ is a generalized Kronecker delta. Therefore

$$K=\mathfrak{K}/\gamma=\sqrt{(G^2-G_{\sigma\rho}G^{\sigma\rho})}, \quad (3.4)$$

which is of the form (3.11). Other relative invariants may be dealt with in an exactly similar way.

4. Consider a small variation of K in which the $g^{\mu\nu}$ are not varied. From (3.1) we have, since $\Gamma^{\mu_1\nu_1\dots\rho_m}$ involves only the $g^{\mu\nu}$,

$$\delta K = \sum_{r=1}^{m} \Omega^{\mu\nu\sigma\rho}_{(r)}\,\delta B_{\mu\nu\sigma\rho}, \tag{4.1}$$

where
$$\Omega^{\mu\nu\sigma\rho}_{(r)} = \Gamma^{\mu_1\nu_1\dots\rho_{r-1}\mu\nu\sigma\rho\mu_{r+1}\dots\rho_m} \prod_{s=1}^{m}{}' B_{\mu_s\nu_s\sigma_s\rho_s}, \tag{4.2}$$

the dash indicating that, in forming the product, the term $s = r$ is to be omitted. But

$$B_{\mu\nu\sigma\rho} = \tfrac{1}{2}(g_{\mu\nu,\sigma\rho}+g_{\sigma\rho,\mu\nu}-g_{\mu\sigma,\nu\rho}-g_{\nu\rho,\mu\sigma})+\{\mu\nu,\alpha\}[\sigma\rho,\alpha]-\{\mu\sigma,\alpha\}[\nu\rho,\alpha], \tag{4.3}$$

so that it is at once apparent from (4.1) that *all* the terms of $\partial K/\partial g_{\mu\nu,\alpha}$ have three-index symbols as factors. Hence, by (2.8), all the terms of $Y^{\mu\nu\alpha}$ have three-index symbols as factors. But $Y^{\mu\nu\alpha}$ is a tensor; and therefore it must be a zero tensor. Hence

$$\frac{\partial K}{\partial g_{\mu\nu,\alpha}} = -2\{\sigma\rho,\mu\}Z^{\nu\sigma\rho\alpha}-2\{\sigma\rho,\nu\}Z^{\mu\sigma\rho\alpha}-\{\sigma\rho,\alpha\}Z^{\mu\nu\sigma\rho}. \tag{4.4}$$

The coefficient of θ'' is accordingly simply

$$K_{;\alpha\beta}\,Z^{\mu\nu\alpha\beta}+2K_{;\alpha}\,Z^{\mu\nu\alpha\beta}{}_{;\beta}. \tag{4.5}$$

5. In re-expressing $P^{\mu\nu}$ as given by (2.3) it is unnecessary to write down explicitly those terms which are multiplied by three-index symbols, for such terms cannot by themselves form the components of a tensor, [(**1**) 79] and we shall therefore represent them merely by dots. (This procedure is equivalent to using geodesic coordinates.)

Now, by (2.61) and (2.62), we have

$$\left.\begin{aligned}\gamma_{,\beta} &= \dots\\ \gamma_{,\alpha\beta} &= \gamma\{\beta\lambda,\lambda\}_{,\alpha}+\dots,\end{aligned}\right\} \tag{5.1}$$

and

$$Z^{\mu\nu\alpha\beta}{}_{,\alpha\beta} = Z^{\mu\nu\alpha\beta}{}_{,\beta\alpha} = (Z^{\mu\nu\alpha\beta}{}_{;\beta})_{,\alpha}-\{\sigma\beta,\mu\}_{,\alpha}\,Z^{\sigma\nu\alpha\beta}-\{\sigma\beta,\nu\}_{,\alpha}\,Z^{\mu\sigma\alpha\beta}-$$
$$-\{\sigma\beta,\alpha\}_{,\alpha}\,Z^{\mu\nu\sigma\beta}-\{\sigma\beta,\beta\}_{,\alpha}\,Z^{\mu\nu\alpha\sigma}+\dots, \tag{5.21}$$
$$(Z^{\mu\nu\alpha\beta}{}_{;\beta})_{,\alpha} = Z^{\mu\nu\alpha\beta}{}_{;\alpha\beta}+\dots.$$

Also write $[\partial K/\partial g_{\mu\nu}]$ for the 'tensor part' of $\partial K/\partial g_{\mu\nu}$, i.e. for those terms of $\partial K/\partial g_{\mu\nu}$ which form the components of a tensor, so that

$$\frac{\partial K}{\partial g_{\mu\nu}} = \left[\frac{\partial K}{\partial g_{\mu\nu}}\right]+\dots. \tag{5.22}$$

Using (2.4), (2.41) and substituting (5.1), (5.21) and (5.22) in (2.3) we reduce the latter to

$$P^{\mu\nu} = \tfrac{1}{2}g^{\mu\nu}K + \left[\frac{\partial K}{\partial g_{\mu\nu}}\right] + Z^{\mu\nu\alpha\beta}{}_{;\alpha\beta} + \{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} + \{\sigma\rho,\nu\}_{,\alpha}\, Z^{\mu\sigma\rho\alpha} + \dots. \tag{5.3}$$

6. From (3.1) we see, as in § 7, that $Z^{\mu\nu\alpha\beta}$ arises from an expression of the form

$$z^{\mu\nu\sigma\rho}(\delta g_{\mu\nu,\sigma\rho} + \delta g_{\sigma\rho,\mu\nu} - \delta g_{\mu\sigma,\nu\rho} - \delta g_{\nu\rho,\mu\sigma}), \tag{6.1}$$

where $z^{\mu\nu\sigma\rho}$ is a tensor.

On changing dummy indices (6.1) becomes

$$(z^{\mu\nu\sigma\rho} + z^{\sigma\rho\mu\nu} - z^{\mu\sigma\nu\rho} - z^{\rho\nu\sigma\mu})\,\delta g_{\mu\nu,\sigma\rho} = \zeta^{\mu\nu\sigma\rho}\,\delta g_{\mu\nu,\sigma\rho} \quad \text{(say)}. \tag{6.11}$$

$Z^{\mu\nu\sigma\rho}$ is therefore the part of $\zeta^{\mu\nu\sigma\rho}$ which is symmetrical in μ, ν and in σ, ρ, i.e.

$$Z^{\mu\nu\sigma\rho} = \tfrac{1}{4}(\zeta^{\mu\nu\sigma\rho} + \zeta^{\nu\mu\sigma\rho} + \zeta^{\mu\nu\rho\sigma} + \zeta^{\nu\mu\rho\sigma}). \tag{6.2}$$

From (6.11) and (6.2) it may be verified that, in addition to (2.41), $Z^{\mu\nu\sigma\rho}$ satisfies the cyclic identity

$$Z^{\mu\nu\sigma\rho} + Z^{\mu\sigma\rho\nu} + Z^{\mu\rho\nu\sigma} = 0. \tag{6.3}$$

Hence $$\{\sigma\rho,\mu\}_{,\alpha}(Z^{\nu\sigma\rho\alpha} + Z^{\nu\rho\sigma\alpha} + Z^{\nu\alpha\sigma\rho}) = 0,$$

i.e., by (2.41),

$$\begin{aligned} 2\{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} &= -Z^{\nu\alpha\sigma\rho}\{\sigma\rho,\mu\}_{,\alpha} \\ &= -Z^{\nu\alpha\sigma\rho}(B_{\sigma\alpha\rho.}{}^{\mu} + \{\sigma\alpha,\mu\}_{,\rho}) + \dots \\ &= +Z^{\nu\alpha\sigma\rho}B^{\mu}_{.\,\alpha\rho\sigma} - Z^{\nu\rho\sigma\alpha}\{\sigma\rho,\mu\}_{,\alpha} + \dots \\ &= Z^{\nu\alpha\sigma\rho}B^{\mu}_{.\,\alpha\sigma\rho} - Z^{\nu\sigma\rho\alpha}\{\sigma\rho,\mu\}_{,\alpha} + \dots. \end{aligned}$$

Hence $$\{\sigma\rho,\mu\}_{,\alpha}\, Z^{\nu\sigma\rho\alpha} = \tfrac{1}{3}Z^{\nu\alpha\sigma\rho}B^{\mu}_{.\,\alpha\sigma\rho} + \dots. \tag{6.31}$$

Since $P^{\mu\nu}$ is a tensor, we have therefore from (5.3) that

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} + \left[\frac{\partial K}{\partial g_{\mu\nu}}\right] + \tfrac{1}{2}g^{\mu\nu}K + \tfrac{1}{3}(Z^{\nu\alpha\sigma\rho}B^{\mu}_{.\,\alpha\sigma\rho} + Z^{\mu\alpha\sigma\rho}B^{\nu}_{.\,\alpha\sigma\rho}). \tag{6.4}$$

The last part of (6.4) may be rewritten in a useful form as follows. Consider the tensor $t^{\mu}_{\lambda} = Z^{\mu\nu\sigma\rho}B_{\lambda\nu\sigma\rho}$. By (6.2) this may be written, after some changes in dummy indices, as

$$\begin{aligned} t^{\mu}_{\lambda} = \tfrac{1}{4}[(z^{\mu\nu\sigma\rho} + z^{\sigma\rho\mu\nu})(B_{\lambda\nu\sigma\rho} - B_{\lambda\sigma\nu\rho} + B_{\lambda\nu\rho\sigma} - B_{\lambda\sigma\rho\nu}) + {}& \\ + (z^{\nu\mu\sigma\rho} + z^{\sigma\rho\nu\mu})(B_{\lambda\nu\sigma\rho} - B_{\lambda\rho\nu\sigma} + B_{\lambda\nu\rho\sigma} - B_{\lambda\rho\sigma\nu})], \end{aligned}$$

from which, in view of the symmetry properties of the curvature tensor, and after some further changes in dummy indices, we find that

$$t^{\mu}_{\lambda} = \tfrac{3}{4} B_{\lambda\nu\sigma\rho}(z^{\mu\nu\sigma\rho}+z^{\nu\mu\rho\sigma}+z^{\sigma\rho\mu\nu}+z^{\rho\sigma\nu\mu}). \tag{6.5}$$

Therefore

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho}+\left[\frac{\partial K}{\partial g_{\mu\nu}}\right]+\tfrac{1}{2}g^{\mu\nu}K+$$
$$+\tfrac{1}{4}\{B^{\nu}{}_{.\alpha\sigma\rho}(z^{\mu\alpha\sigma\rho}+z^{\alpha\mu\rho\sigma}+z^{\sigma\rho\mu\alpha}+z^{\rho\sigma\alpha\mu})+$$
$$+B^{\mu}{}_{.\alpha\sigma\rho}(z^{\nu\alpha\sigma\rho}+z^{\alpha\nu\rho\sigma}+z^{\sigma\rho\nu\alpha}+z^{\rho\sigma\alpha\nu})\}. \tag{6.6}$$

7(*a*). If we examine (3.1), we may think of K written formally as the product of m of the components of the curvature tensor with all the indices paired so that to every superscript there corresponds one and only one subscript. Then corresponding to each such pair of indices there is one fact or $g^{\sigma\rho}$ in $\Gamma^{\mu_1\nu_1\ldots\rho_m}$. In forming δK this 'contributes' $-g^{\sigma\mu}g^{\rho\nu}\,\delta g_{\mu\nu}$. So that, if in turn each superscript is replaced by μ, whilst the corresponding subscript is first raised and then replaced by ν, and all $2m$ tensors so obtained are added, then the symmetrical part of the sum is $-[\partial K/\partial g_{\mu\nu}]$.

(*b*) Comparing (4.1) and (6.1) we see at once that

$$z^{\mu\nu\sigma\rho} = \tfrac{1}{2}\sum_{r=1}^{m}\Omega^{\mu\nu\sigma\rho}_{(r)}. \tag{7.1}$$

Hence the first four of the last eight terms of (6.6), say $q^{\mu\nu}$, are

$$q^{\mu\nu} = \tfrac{1}{8}\sum_{r=1}^{m}(\Omega^{\mu\alpha\sigma\rho}_{(r)}B^{\nu}{}_{.\alpha\sigma\rho}+\Omega^{\alpha\mu\sigma\rho}_{(r)}B_{\alpha.\sigma\rho}^{\;\;\nu}+\Omega^{\sigma\rho\mu\alpha}_{(r)}B_{\sigma\rho.\alpha}^{\;\;\;\;\nu}+\Omega^{\sigma\rho\alpha\mu}_{(r)}B_{\sigma\rho\alpha.}^{\;\;\;\;\;\;\nu}). \tag{7.2}$$

Inspection of (7.2) shows at once that $-4q^{\mu\nu}$ is formed by exactly the same process as that described in § 7(*a*), if we notice that each component of the curvature tensor is now taken twice over, which cancels the factor $\tfrac{1}{2}$ on the right-hand side of (7.1). Now, where a superscript was replaced by μ the first time, it is replaced by ν the second time. It follows that $q^{\mu\nu}$ is symmetrical, so that

$$q^{\mu\nu} = q^{\nu\mu} = -\tfrac{1}{4}\left[\frac{\partial K}{\partial g_{\mu\nu}}\right] = +\tfrac{1}{3}Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{.\alpha\sigma\rho}. \tag{7.3}$$

Hence (6.4) becomes

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho}-\tfrac{2}{3}Z^{\nu\alpha\sigma\rho}B^{\mu}{}_{.\alpha\sigma\rho}+\tfrac{1}{2}g^{\mu\nu}K, \tag{7.4}$$

which establishes (1.1) for the case in which K has the form (3.1).

A simple formula may be derived in this case for the scalar invariant $P = g_{\mu\nu} P^{\mu\nu}$. From the law of formation of $[\partial K/\partial g_{\mu\nu}]$ given in § 7(a) it follows at once that

$$g_{\mu\nu}\left[\frac{\partial K}{\partial g_{\mu\nu}}\right] = -2mK. \tag{7.5}$$

Hence, with (7.3), (7.4) easily yields

$$P = (g_{\mu\nu} Z^{\mu\nu\sigma\rho})_{;\sigma\rho} + (\tfrac{1}{2}n - m)K. \tag{7.6}$$

When $m = \frac{1}{2}n$, we see that P has necessarily the form of a divergence.

8. I now show that (7.4) holds in the general case when K has the form (3.11). Write

$$P_i^{\mu\nu} = \frac{hK_i}{hg_{\mu\nu}}, \quad \text{and} \quad f_i = \frac{\partial F}{\partial K_i}, \qquad f_{ij} = \frac{\partial^2 F}{\partial K_i \partial K_j}, \quad \text{etc.}$$

Also let

$$Z_i^{\mu\nu\sigma\rho} = \frac{\partial K_i}{\partial g_{\mu\nu,\sigma\rho}}.$$

Then using the summation convention also for roman indices it is not difficult to show after the manner of §§ 2, 4 that

$$P^{\mu\nu} = \frac{hK}{hg_{\mu\nu}} = \tfrac{1}{2}g^{\mu\nu}F + f_i(P_i^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}K_i) +$$
$$+ f_{ij}[(K_j)_{;\sigma\rho} Z_i^{\mu\nu\sigma\rho} + 2(K_j)_{;\sigma} Z_i^{\mu\nu\sigma\rho}{}_{;\rho}] + f_{ijk}(K_j)_{;\sigma}(K_k)_{;\rho} Z_i^{\mu\nu\sigma\rho}. \tag{8.1}$$

Now consider (7.4) with $K = F$. We have

$$Z^{\mu\nu\sigma\rho} = f_i Z_i^{\mu\nu\sigma\rho}.$$

Therefore

$$P^{\mu\nu} = (f_i Z_i^{\mu\nu\sigma\rho})_{;\sigma\rho} - \tfrac{2}{3} B^{\mu}{}_{.\alpha\sigma\rho} f_i Z_i^{\nu\alpha\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}F. \tag{8.2}$$

But $\qquad (f_i Z_i^{\mu\nu\sigma\rho})_{;\sigma\rho} = f_i Z_i^{\mu\nu\sigma\rho}{}_{;\sigma} + f_{ij}(K_j)_{;\sigma} Z_i^{\mu\nu\sigma\rho}, \quad$ etc.

Hence comparing (8.2) and (8.1) we see that (7.4) applies when K has the general form (3.11). This completes the proof.

9. It is scarcely worth while verifying in full detail that the divergence of (7.4) vanishes identically. I shall merely outline a proof for the case in which K is of the form (3.1). We have

$$P^{\mu\nu}{}_{;\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} - \tfrac{2}{3}(Z^{\nu\alpha\sigma\rho} B^{\mu}{}_{.\alpha\sigma\rho})_{;\nu} + \tfrac{1}{2}K^{;\mu}. \tag{9.1}$$

If we add to $Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu}$ the expressions $Z^{\mu\nu\sigma\rho}{}_{;\sigma\nu\rho}$ and $Z^{\mu\nu\sigma\rho}{}_{;\nu\sigma\rho}$, both of which may be re-expressed in terms of $Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu}$ by means of the Ricci's identity, we obtain

$$3Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} = 2B^{\mu}{}_{.\rho\nu\alpha} Z^{\alpha\nu\sigma\rho}{}_{;\sigma}, \tag{9.2}$$

A 825

all other terms vanishing on account of the symmetry properties of $Z^{\mu\nu\sigma\rho}$ and $B_{\mu\nu\sigma\rho}$, and of $q^{\mu\nu}$. For the same reason, when we substitute (9.2) in (9.1), the latter can be reduced to

$$P^{\mu\nu}{}_{;\nu} = -\tfrac{2}{3}B^{\mu}{}_{.\alpha\sigma\rho;\nu}\,Z^{\nu\alpha\sigma\rho}+\tfrac{1}{2}K^{;\mu}. \tag{9.3}$$

From (3.1), (4.1) and (7.1) we have at once that

$$\tfrac{1}{2}K^{;\mu} = z^{\beta\alpha\sigma\rho}B_{\beta\alpha\sigma\rho}{}^{;\mu}. \tag{9.4}$$

If the first term of (9.3) be now rewritten in terms of $z^{\mu\nu\sigma\rho}$, which can be done by inspection of (6.4) and (6.6), it is not difficult to show by applying the Bianchi identity to the resulting expression that it just cancels the second term, so that

$$P^{\mu\nu}{}_{;\nu} \equiv 0. \tag{9.5}$$

10. I shall now consider a few examples of specific invariants K.

(i)
$$K = G \quad (= g^{\mu\nu}g^{\sigma\rho}B_{\mu\nu\sigma\rho}). \tag{10.1}$$

By inspection $z^{\mu\nu\sigma\rho} = \tfrac{1}{2}g^{\mu\nu}g^{\sigma\rho}$, giving

$$Z^{\mu\nu\sigma\rho} = (g^{\mu\nu}g^{\sigma\rho}-\tfrac{1}{2}g^{\mu\sigma}g^{\nu\rho}-\tfrac{1}{2}g^{\mu\rho}g^{\nu\sigma}).$$

Therefore, at once, since here $Z^{\mu\nu\sigma\rho}$ vanishes,

$$\frac{hG}{hg_{\mu\nu}} = -G^{\mu\nu}+\tfrac{1}{2}g^{\mu\nu}G. \tag{10.2}$$

(ii)
$$K = G^2 = K''' \quad \text{(say)}. \tag{10.3}$$

It is easiest to use (8.2) with $F = K_1^2$; $K_1 = G$. It follows immediately that

$$P'''^{\mu\nu} = \frac{hG^2}{hg_{\mu\nu}} = -2\{G^{;\mu\nu}+GG^{\mu\nu}-g^{\mu\nu}\quad(\Box G+\tfrac{1}{4}G^2)\}. \tag{10.4}$$

[The symbol $\Box$ before any expression (...) represents the operation $g^{\sigma\rho}(\ldots)_{;\sigma\rho}$].

(iii)
$$K = G_{\mu\nu}G^{\mu\nu} = K' \quad \text{(say)}. \tag{10.5}$$

For a small variation of the $g_{\mu\nu,\sigma\rho}$ alone

$$\delta K' = 2G^{\mu\nu}\delta G_{\mu\nu} = 2G^{\mu\nu}g^{\sigma\rho}\delta B_{\mu\nu\sigma\rho},$$

so that $z^{\mu\nu\sigma\rho} = g^{\sigma\rho}G^{\mu\nu}$, and

$$Z^{\mu\nu\sigma\rho} = g^{\mu\nu}G^{\sigma\rho}+g^{\sigma\rho}G^{\mu\nu}-\tfrac{1}{2}g^{\mu\sigma}G^{\nu\rho}-\tfrac{1}{2}g^{\nu\sigma}G^{\mu\rho}-\tfrac{1}{2}g^{\mu\rho}G^{\nu\sigma}-\tfrac{1}{2}g^{\nu\rho}G^{\mu\sigma}.$$

Using the identities of Ricci and Bianchi we find

$$Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} = \Box G^{\mu\nu}-G^{;\mu\nu}-B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot}G_{\sigma\rho}+G^{\mu\rho}G^{\nu}_{\rho}+\tfrac{1}{2}g^{\mu\nu}\Box G.$$

Also
$$Z^{\mu\alpha\sigma\rho}B^{\nu}{}_{.\alpha\sigma\rho} = \tfrac{3}{2}(B^{\mu\nu\sigma\rho}_{\cdot\cdot\cdot\cdot}G_{\sigma\rho}+G^{\mu\rho}G^{\nu}_{\rho}).$$

Therefore

$$P'^{\mu\nu} = \frac{hK'}{hg_{\mu\nu}} = \Box G^{\mu\nu} - G^{;\mu\nu} - 2B^{\mu\nu\sigma\rho}_{\cdots\cdot}G_{\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}(\Box G + G_{\sigma\rho}G^{\sigma\rho}). \tag{10.6}$$

(iv) $$K = B_{\mu\nu\sigma\rho}B^{\mu\nu\sigma\rho}_{\cdots\cdot} = K'' \quad \text{(say)}. \tag{10.7}$$

By inspection, $z^{\mu\nu\sigma\rho} = B^{\mu\nu\sigma\rho}$, which, in view of the symmetry properties of $B_{\mu\nu\sigma\rho}$, gives

$$Z^{\mu\nu\sigma\rho} = 2(B^{\mu\nu\sigma\rho}_{\cdots\cdot} + B^{\mu\nu\rho\sigma}_{\cdots\cdot}).$$

Hence $$Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} = 4\Box G^{\mu\nu} - 2G^{;\mu\nu} - 4B^{\mu\nu\sigma\rho}_{\cdots\cdot}G_{\sigma\rho} + 4G^{\mu\sigma}G^{\nu}_{\sigma}.$$

Again, in view of the symmetry properties of the curvature tensor

$$Z^{\mu\alpha\sigma\rho}B^{\nu}_{\cdot\alpha\sigma\rho} = 2(B^{\mu\alpha\sigma\rho}_{\cdots\cdot}B^{\nu}_{\cdot\alpha\sigma\rho} + B^{\mu\alpha\sigma\rho}_{\cdots\cdot}B^{\nu}_{\cdot\alpha\rho\sigma}) = 3B^{\mu\alpha\sigma\rho}_{\cdots\cdot}B^{\nu}_{\cdot\alpha\sigma\rho}.$$

Therefore

$$\begin{aligned} P''^{\mu\nu} &= \frac{hK''}{hg_{\mu\nu}} \\ &= 4\{\Box G^{\mu\nu} - \tfrac{1}{2}G^{;\mu\nu} - B^{\mu\nu\sigma\rho}_{\cdots\cdot}G_{\sigma\rho} + G^{\mu\sigma}G^{\nu}_{\sigma} - \tfrac{1}{2}B^{\mu\alpha\sigma\rho}_{\cdots\cdot}B^{\nu}_{\cdot\alpha\sigma\rho} + \tfrac{1}{8}g^{\mu\nu}K''\}. \end{aligned} \tag{10.8}$$

11. In the case of the three invariants K', K'', K''', we have $m = 2$ so that, according to (7.6), their contracted Hamiltonian derivatives P', P'', P''' must each have the form of a divergence in a V_4. In fact we find

$$\tfrac{1}{2}P' = \tfrac{1}{2}P'' = \tfrac{1}{6}P''' = \Box G. \tag{11.1}$$

It will be noticed that the terms involving the fourth derivatives of the $g_{\mu\nu}$ may be eliminated between the three Hamiltonian derivatives (10.4), (10.6), and (10.8). Thus, if we write

$$\left.\begin{aligned} L^{\mu\nu} &= B^{\mu\alpha\sigma\rho}_{\cdots\cdot}B^{\nu}_{\cdot\alpha\sigma\rho} + GG^{\mu\nu} - 2G^{\mu\sigma}G^{\nu}_{\sigma} - 2B^{\mu\nu\sigma\rho}_{\cdots\cdot}G_{\sigma\rho} \\ L &= g_{\mu\nu}L^{\mu\nu} = -4K' + K'' + K''' \end{aligned}\right\}, \tag{11.2}$$

we have $$-\frac{1}{2}\frac{hL}{hg_{\mu\nu}} = L^{\mu\nu} - \tfrac{1}{4}g^{\mu\nu}L. \tag{11.3}$$

Accordingly there exists in a V_n ($n > 4$, see below) a linear combination with constant coefficients of the three invariants K', K'', K''', namely L, such that its Hamiltonian derivative is a tensor of the *second* differential order. This naturally implies a considerable restriction on any common integrals which the three sets of differential equations

$$P'^{\mu\nu} = 0, \qquad P''^{\mu\nu} = 0, \qquad P'''^{\mu\nu} = 0 \tag{11.4}$$

A 826

may possess, for they must satisfy the set of *second*-order equations

$$L^{\mu\nu} = 0. \tag{11.5}$$

It has been shown by Lanczos (2) that $hL/hg_{\mu\nu}$ vanishes identically in a V_4. The same result holds in a V_2 and a V_3.

REFERENCES

1. Eddington, *The Mathematical Theory of Relativity* (2nd edn. Cambridge, 1930).
2. Lanczos, *Annals of Math.* (2) 39 (1938), 842.
3. Veblen, *Invariants of Quadratic Differential Forms* (Cambridge Tract 24, 1927).

287

It has been published as: Buchdahl, H. A., 1948. A special class of solutions of the equations of the gravitational field arising from certain gauge-invariant action principles, Proceedings of the National Academy of Sciences, 34(2), 66-68

ÜBER DIE VARIATIONSABLEITUNG VON FUNDAMENTALINVARIANTEN BELIEBIG HOHER ORDNUNG.

von H.A. Buchdahl, Physics Department,
University of Tasmania.

§1. Es sei K eine Fundamentalinvariante, in welcher die Ableitungen der Komponenten des metrischen Tensors $g_{\mu\nu}$ bis zur ($\underline{s+2}$)-ten Ordnung vorkommen, wobei $\underline{s}$ irgendeine positive ganze Zahl oder Null ist. Dann ist die Variationsableitung $P^{\mu\nu}$ von K durch die Gleichung

$$\delta\int \underline{K} d\tau \equiv \delta\int K\sqrt{-g}\, d\tau = \int P^{\mu\nu}\delta g_{\mu\nu}\sqrt{-g}\, d\tau \qquad (1.1)$$

definiert, wobei vorausgesetzt wird, dass alle Variationen der $g_{\mu\nu}$ und ihrer Ableitungen an der Grenze des betrachteten Integrationsgebietes verschwinden. $P^{\mu\nu}$ ist selbstverständlich ein kontravarianter symmetrischer Fundamentaltensor, der die Eigenschaft besitzt, dass seine Divergenz identisch verschwindet. (Eddington, 1930, §61).

Indem wir die Schreibweise von Schouten (Schouten, 1924, S. 25 und 31) für den symmetrischen oder den alternierenden Teil eines Tensors verwenden, wollen wir nun zeigen, dass sich die Variationsableitung $P^{\mu\nu}$ in einer einfachen Form darstellen lässt, die einerseits ihren Tensorcharakter klar erkennen lässt, und die sich andererseits zu ihrer Berechnung eignet, wenn K explizit gegeben ist. Das zu beweisende Ergebnis lässt sich dann in die folgende Gestalt bringen

$$P^{\mu\nu} = \tfrac{1}{2}X^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{1}{3}B^{(\mu}{}_{\cdot\alpha\sigma\rho}X^{\nu)\alpha\sigma\rho} + \tfrac{1}{2}g^{\mu\nu}K + (T^{\varepsilon(\mu\nu)} - T^{(\mu\nu)\varepsilon})_{;\varepsilon} + t^{(\mu\nu)}. \qquad (1.2)$$

Dabei sind die Tensoren $X^{\mu\nu\sigma\varrho}$, $T^{\alpha\beta\gamma}$, $t^{\mu\nu}$ aufgebaut aus dem Krümmungstensor $B_{\mu\nu\sigma\varrho}$ und seinen kovarianten Ableitungen, und den Tensoren

$$E^{\sigma_1\cdots\sigma_j} = \frac{\partial K}{\partial B_{\sigma_1\cdots\sigma_4;\sigma_5\cdots\sigma_j}} \qquad (1.3)$$

und ihren kovarianten Ableitungen.

Steht ein unterer Index hinter einem Komma, so bedeutet er gewöhnliche Differentiation, steht er hinter einem Semikolon, so bedeutet er kovariante Differentiation. Die Symmetrieeigenschaften des Krümmungstensors sind auf der rechten Seite von (1.3) unbeachtet zu lassen, die Reihenfolge seiner Indizes ist die von Eddington benutzte, (Eddington, 1930, S. 72). Definieren wir

$$F^{\sigma_1\cdots\sigma_j} = E^{\sigma_1\cdots\sigma_j} + \sum_{r=1}^{s-j+4}(-1)^r E^{\sigma_1\cdots\sigma_{j+r}}{}_{;\sigma_{j+r}\cdots\sigma_{j+1}}, \qquad (1.4)$$

so können wir etwas ausführlicher schreiben:

$$\left.\begin{aligned} X^{\mu\nu\sigma\varrho} &= \bar{X}^{(\mu\nu)(\sigma\varrho)}, \\ \text{mit} \quad \bar{X}^{\mu\nu\sigma\varrho} &= 4F^{[\mu[\nu\sigma]\varrho]}; \end{aligned}\right\} \qquad (1.5)$$

und

$$T^{\alpha\beta\gamma} = \frac{1}{2}\sum_{j=0}^{s-1}\sum_{i=1}^{j+4} F^{\sigma_1\cdots\sigma_{i-1}\alpha\sigma_{i+1}\cdots\sigma_{j+4}\beta} B_{\sigma_1\cdots;\cdots\sigma_{i-1}}{}^{\gamma}{}_{\sigma_{i+1}\cdots\sigma_{j+4}}, \qquad (1.6)$$

$$t^{\mu\nu} = -\frac{1}{2}\sum_{j=0}^{s-1} F^{\sigma_1\cdots\sigma_{j+4}\mu} B_{\sigma_1\cdots;\cdots\sigma_{j+4}}{}^{\nu}. \qquad (1.7)$$

Es lässt sich sodann durch direkte Rechnung nachweisen, dass die Divergenz von $P^{\mu\nu}$ tatsächlich identisch verschwindet, (§ 6). In §§ 7 und 8 betrachten wir einige spezielle Ergebnisse.

Die Dimensionszahl des betrachteten Raumes ist durchwegs als n angenommen.

§ 2. Schreiben wir der Einfachheit halber γ für $\sqrt{-g}$, so haben wir

$$\delta J = \delta\int K\gamma\, d\tau = \int\left(\tfrac{1}{2}g^{\mu\nu}K\delta g_{\mu\nu} + \delta K\right)\gamma\, d\tau \quad . \qquad (2.1)$$

(Sollte es vorkommen, dass uns statt K eine Tensordichte $\underline{K}$ gegeben ist, die mit Hilfe der numerischen Tensordichten $\varepsilon^{\mu_1\cdots\mu_n}$

und $\varepsilon_{\mu_1\cdots\mu_n}$ von Levi-Civitā gebildet ist, (Veblen, 1927, S. 25), so bringen wir sie erst in die Form KY, was unter Zuhilfenahme der Identitäten

$$\left.\begin{aligned} \varepsilon^{\mu_1\cdots\mu_n} g_{\mu_1\nu_1} g_{\mu_2\nu_2}\cdots g_{\mu_n\nu_n} &= g\varepsilon_{\nu_1\cdots\nu_n} , \\ \varepsilon^{\mu_1\cdots\mu_n}\varepsilon_{\nu_1\cdots\nu_n} &= \delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n} \end{aligned}\right\} \qquad (2.2)$$

immer möglich ist, wobei man dann $\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}$ als eine Determinante der gewöhnlichen Kroneckerschen Symbole $\delta^{\mu_i}_{\nu_k}$ betrachtet).

Jede Invariante K ist nun einfach eine Funktion einer gewissen Anzahl von Invarianten die formal allein ein Produkt von Komponenten des kontravarianten metrischen Tensors, des Krümmungstensors und dessen kovarianten Ableitungen sind. Wir schreiben

$$K_{(p)} = \Gamma C \qquad (2.3)$$

für die p-te dieser Invarianten. Wir haben dabei die Indizes der Tensoren Γ und C unterdrückt. Γ ist einfach ein Produkt kontravarianter Komponenten des metrischen Tensors, während C ein Produkt von m_i Komponenten der i-ten $(i=0,1,\ldots,s)$ kovarianten Ableitungen des kovarianten Krümmungstensors ist. Wir sagen $K_{(p)}$ sei von der Ordnung $s+2$ und vom Grade

$$M_{(p)} = \sum_{i=0}^{s} m_i . \qquad (2.4)$$

Γ ist offensichtlich $\overline{M}_{(p)}$-ter Stufe, wobei

$$\overline{M}_{(p)} = \sum_{i=0}^{s} (i+4)m_i . \qquad (2.41)$$

Unterdrückt man nun in $K_{(p)}$ der Reihe nach jeden einzelnen der darin vorkommenden m_i Faktoren $B_{\mu_{1ij}\cdots j\cdots\mu_{i+1j}}$ $(j=1,2,\ldots m_i)$, und ersetzt jedesmal die so freigewordenen Indizes in der richtigen Reihenfolge durch $\sigma_1,\ldots\sigma_{i+4}$, so erhält man m_i Tensoren $(i+4)$-ter Stufe, deren Summe wir mit $E_{(p)}^{\sigma_1\cdots\sigma_{i+4}}$ bezeichnen. Man hat also

$$E_{(p)}^{\sigma_1\cdots\sigma_j} = \frac{\partial K}{\partial B_{\sigma_1\cdots j\cdots\sigma_j}} , \qquad (2.5)$$

wobei irgendwelche Symmetrieeigenschaften der $B_{\sigma_1 \dots ; \dots \sigma_j}$ ausser acht zu lassen sind. Weiterhin schreiben wir

$$\eta_{(p)}^{\mu\nu} = C \frac{\partial \Gamma}{\partial g_{\mu\nu}} \,. \tag{2.51}$$

Da nun K von der Form

$$K = F(K_{(1)}, K_{(2)}, \dots, K_{(p)}, \dots) \tag{2.6}$$

ist, haben wir sofort

$$\delta K = \eta^{\mu\nu} \delta g_{\mu\nu} + \sum_{j=0}^{s} E^{\sigma_1 \cdots \sigma_{j+4}} \delta B_{\sigma_1 \dots ; \dots \sigma_{j+4}} \,, \tag{2.7}$$

wobei

$$\eta^{\mu\nu} = \sum_p \eta_{(p)}^{\mu\nu} \frac{\partial F}{\partial K_{(p)}} \tag{2.71}$$

und

$$E^{\sigma_1 \cdots \sigma_j} = \sum_p E_{(p)}^{\sigma_1 \cdots \sigma_j} \frac{\partial F}{\partial K_{(p)}} = \frac{\partial K}{\partial B_{\sigma_1 \dots ; \dots \sigma_j}} \tag{2.72}$$

ist. Nun kann man sich $K_{(p)}$ einfach als das Produkt von $M_{(p)}$ Faktoren $B^{\cdots}_{\cdots ; \cdots}$ geschrieben denken, nämlich so, dass jedem oberen Index genau ein gleicher unterer Index entspricht. Wir bezeichnen die Summe der $\frac{1}{2}\overline{M}_{(p)}$ Tensoren die man erhält ~~die man erhelt~~, indem man der Reihe nach jeden oberen Index durch μ ersetzt und gleichzeitig den ihm entsprechenden unteren Index heraufzieht und durch ν ersetzt, mit $\tilde{\eta}_{(p)}^{\mu\nu}$. Nun ist aber $\delta g^{\mu\nu} = - g^{\mu\sigma} g^{\nu\varrho} \delta g_{\sigma\varrho}$, daher folgt sofort

$$\eta_{(p)}^{\mu\nu} = - \tilde{\eta}_{(p)}^{(\mu\nu)} \,. \tag{2.8}$$

Wir haben also, wegen (2.71), (2.72),

$$\eta^{\mu\nu} = -\frac{1}{2} \sum_{j=0}^{s} \sum_{i=1}^{j+4} E^{\sigma_1 \cdots \sigma_{i-1} \mu \sigma_{i+1} \cdots \sigma_{j+4}} B_{\sigma_1 \dots ; \dots \sigma_{i-1}}{}^{\nu}{}_{\sigma_{i+1} \cdots \sigma_{j+4}} \,. \tag{2.9}$$

Wir betonen ausdrücklich noch einmal, dass die rechte Seite von (2.9) an sich schon in μ und ν symmetrisch ist,

$$\eta^{[\mu\nu]} = 0 \,, \tag{2.91}$$

da dabei jeder Index von $K_{(p)}$ zweimal ersetzt wird, einmal durch

μ, das andere mal durch ν, sodass die Doppelsumme einfach $\tilde{\eta}^{\mu\nu} + \tilde{\eta}^{\nu\mu}$ ist. (Die Bedutung der $\tilde{\eta}^{\mu\nu}$ versteht sich von selbst). Wir erhalten daher aus (2.1)

$$\delta J = \int [(\tfrac{1}{2} g^{\mu\nu} K + \eta^{\mu\nu}) \delta g_{\mu\nu} + \sum_{j=0}^{s} E^{\sigma_1 \cdots \sigma_{j+\nu}} \delta B_{\sigma_1 \cdots j \cdots \sigma_{j+\nu}}] \gamma \, d\tau \quad . \qquad (2.10)$$

§3. Wir beweisen jetzt zunächst die folgende Beziehung: Sind die Tensonen $A^{\sigma_1 \cdots \sigma_k \rho}$ und $B_{\sigma_1 \cdots \sigma_k}$ kontravariant (<u>k+1</u>)-ter Stufe, bzw. kovariant <u>k</u>-ter Stufe, dann ist

$$\delta I \equiv \int \gamma A^{\sigma_1 \cdots \sigma_k \rho} \delta B_{\sigma_1 \cdots \sigma_k ; \rho} \, d\tau$$

$$= \int \gamma [-A^{\sigma_1 \cdots \sigma_k \rho}{}_{;\rho} \delta B_{\sigma_1 \cdots \sigma_k} + (D^{(\mu|\varepsilon|\nu)} + D^{\varepsilon(\mu\nu)} - D^{(\mu\nu)\varepsilon})_{;\varepsilon} \delta g_{\mu\nu}] d\tau \quad , \qquad (3.1)$$

vorausgesetzt, dass alle Variationen der $g_{\mu\nu}$ und ihrer Ableitungen an der Grenze des betrachteten Integrationsgebietes veschwinden. Hierbei ist

$$D^{\mu\varepsilon\nu} = \frac{1}{2} \sum_{i=1}^{k} A^{\sigma_1 \cdots \sigma_{i-1} \mu \sigma_{i+1} \cdots \sigma_k \varepsilon} B_{\sigma_1 \cdots \sigma_{i-1}}{}^{\nu}{}_{\sigma_{i+1} \cdots \sigma_k} \quad . \qquad (3.11)$$

Wir haben

$$\delta I = \int \gamma A^{\sigma_1 \cdots \sigma_k \rho} \delta (B_{\sigma_1 \cdots \sigma_k , \rho} - \sum_{i=1}^{k} \{\sigma_i \rho, \varepsilon\} B_{\sigma_1 \cdots \sigma_{i-1} \varepsilon \sigma_{i+1} \cdots \sigma_k}) d\tau$$

$$= -\int \gamma [(A^{\sigma_1 \cdots \sigma_k \rho}{}_{,\rho} + \{\rho\lambda, \lambda\} A^{\sigma_1 \cdots \sigma_k \rho}) \delta B_{\sigma_1 \cdots \sigma_k}$$

$$+ \sum_{i=1}^{k} A^{\sigma_1 \cdots \sigma_k \rho} B_{\sigma_1 \cdots \sigma_{i-1} \varepsilon \sigma_{i+1} \cdots \sigma_k} \delta \{\sigma_i \rho, \varepsilon\}$$

$$+ \sum_{i=1}^{k} \{\sigma_i \rho, \varepsilon\} A^{\sigma_1 \cdots \sigma_k \rho} \delta B_{\sigma_1 \cdots \sigma_{i-1} \varepsilon \sigma_{i+1} \cdots \sigma_k}] d\tau$$

$$= -\int \gamma \delta L \, d\tau \quad , \qquad (3.2)$$

da der integrierte Teil laut Voraussetzung verschwindet. Nun ist δL ein Tensor (Skalar). Wir deuten daher all Glieder, welche Christoffelsche Dreizeigersymbole als Faktoren enthalten, durch Punkte an, da solche Glieder allein keinen Tensor bilden können. Für das erste Glied in (3.2) erhalten wir daher

$$\int \gamma A^{\sigma_1 \cdots \sigma_k \rho}{}_{,\rho} \, d\tau = \int \gamma (A^{\sigma_1 \cdots \sigma_k \rho}{}_{;\rho} + \ldots) d\tau \quad . \qquad (3.3)$$

Die erste Summe in (3.2) ergibt

$$\int \sum_{i=1}^{k} \gamma A^{\sigma_1 \cdots \sigma_k \varrho} B_{\sigma_1 \cdots \sigma_{i-1} \varepsilon \sigma_{i+1} \cdots \sigma_k} \delta(g^{\varepsilon\lambda}[\sigma_i \varrho, \lambda]) d\tau$$

$$= \int \gamma \left[\tfrac{1}{2} \sum_{i=1}^{k} A^{\sigma_1 \cdots \sigma_k \varrho} B_{\sigma_1 \cdots \sigma_{i-1}}{}^{\lambda}{}_{\sigma_{i+1} \cdots \sigma_k} \delta(g_{\sigma_i \lambda, \varrho} + g_{\varrho\lambda, \sigma_i} - g_{\sigma_i \varrho, \lambda}) + \ldots \right] d\tau \; ,$$

während andererseits sowohl das zweite Glied als auch die zweite Summe in (3.2) offensichtlich durch Punkte ersetzt werden kann. Durch partielle Integration und Umordnen von Indizes erhält man sodann (3.1) ohne weiteres.

§4. Wendet man jetzt (3.1) wiederholt auf (2.10) an, so ergibt sich sofort $\delta J = \int [(\tfrac{1}{2} g^{\mu\nu} K + \eta^{\mu\nu} + S^{\mu\nu}) \delta g_{\mu\nu} + F^{\mu\nu\sigma\varrho} \delta B_{\mu\nu\sigma\varrho}] \gamma \, d\tau$, (4.1)

mit $$S^{\mu\nu} = (T^{(\mu|\varepsilon|\nu)} + T^{\varepsilon(\mu\nu)} - T^{(\mu\nu)\varepsilon})_{;\varepsilon} \; . \qquad (4.11)$$

$T^{\alpha\beta\gamma}$ ist durch (1.6) definiert, während $F^{\mu\nu\sigma\varrho}$ ein Spezialfall von (1.4) ist, (mit j=4). Nun ist

$$\delta J' \equiv \int \gamma F^{\mu\nu\sigma\varrho} \delta B_{\mu\nu\sigma\varrho} \, d\tau$$

$$= \int \gamma F^{\mu\nu\sigma\varrho} [\tfrac{1}{2}(\delta g_{\mu\nu,\sigma\varrho} + \delta g_{\sigma\varrho,\mu\nu} - \delta g_{\mu\sigma,\nu\varrho} - \delta g_{\nu\varrho,\mu\sigma})$$

$$+ \{\mu\nu, \alpha\} \delta [\varrho\sigma, \alpha] + [\varrho\sigma, \alpha] \delta \{\mu\nu, \alpha\}$$

$$- \{\mu\sigma, \alpha\} \delta [\nu\varrho, \alpha] - [\nu\varrho, \alpha] \delta \{\mu\sigma, \alpha\}] d\tau$$

$$= \tfrac{1}{2} \int \gamma X^{\mu\nu\sigma\varrho} [\delta g_{\mu\nu,\sigma\varrho} + \{\sigma\varrho, \alpha\}(\delta g_{\mu\alpha,\nu} + \delta g_{\nu\alpha,\mu} - \delta g_{\mu\nu,\alpha}) + \ldots] d\tau \; . \qquad (4.2)$$

Hierbei ist $X^{\mu\nu\sigma\varrho}$ durch (1.5) definiert, wahrend die Bedeutung der Punkte in §3 angegeben ist. Wenden wirt jetzt partielle Integration an, und ersetzen weitherhin gewöhnliche durch kovariante Ableitungen, so lässt sich (4.2) in die Form

$$\delta J' = \tfrac{1}{2} \int \gamma [X^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} - (\{\varrho\varepsilon, \nu\}_{,\sigma} + \{\varrho\sigma, \nu\}_{,\varepsilon}) X^{\mu\varepsilon\sigma\varrho}$$

$$- (\{\varrho\varepsilon, \mu\}_{,\sigma} + \{\varrho\sigma\mu\}_{,\varepsilon}) X^{\nu\varepsilon\sigma\varrho} + \ldots] \delta g_{\mu\nu} d\tau \qquad (4.3)$$

bringen. Infolge (1.5) ist es selbstverständlich, dass

$$X^{[\mu\nu]\sigma\varrho} = X^{\mu\nu[\sigma\varrho]} = 0 \qquad (4.4)$$

ist; weiterhin lässt sich die Symmetrieeigenschaft

$$X^{\mu(\nu\sigma\rho)} = 0 \tag{4.41}$$

unschwer nachweisen. Multipliziert man jetzt (4.41) mit $\{\sigma\rho,\nu\}_{,\varepsilon}$, so erhält man

$$\begin{aligned} 2\{\sigma\rho,\nu\}_{,\varepsilon}X^{\mu\sigma\rho\varepsilon} &= -\{\sigma\rho,\nu\}_{,\varepsilon}X^{\mu\varepsilon\sigma\rho} \\ &= (B_{\sigma\rho\varepsilon\cdot}{}^{\nu} - \{\sigma\varepsilon,\nu\}_{,\rho})X^{\mu\varepsilon\sigma\rho} + \ldots \\ &= B_{\sigma\rho\varepsilon\cdot}{}^{\nu}X^{\mu\varepsilon\sigma\rho} - \{\sigma\rho,\nu\}_{,\varepsilon}X^{\mu\rho\sigma\varepsilon} + \ldots , \end{aligned} \tag{4.5}$$

somit $\{\sigma\rho,\nu\}_{,\varepsilon}X^{\mu\sigma\rho\varepsilon} = \frac{1}{3}B^{\nu}{}_{\cdot\varepsilon\sigma\rho}X^{\mu\varepsilon\sigma\rho} + \ldots$. (4.51)

Wenden wir jetzt (4.5) und (4.51) auf (4.3) an, so erhalten wir

$$\delta J' = \frac{1}{2}\int \gamma (X^{\mu\nu\sigma\rho}{}_{;\sigma\rho} + B^{(\mu}{}_{\cdot\alpha\sigma\rho}X^{\nu)\alpha\sigma\rho})\delta g_{\mu\nu}d\tau \; . \tag{4.6}$$

Laut (4.1) und (4.6) nimmt die Variationsableitung von K jetzt die Gestalt

$$P^{\mu\nu} = \frac{1}{2}X^{\mu\nu\sigma\rho}{}_{;\sigma\rho} + \frac{1}{3}B^{(\mu}{}_{\cdot\alpha\sigma\rho}X^{\nu)\alpha\sigma\rho} + S^{\mu\nu} + \eta^{\mu\nu} + \frac{1}{2}g^{\mu\nu}K \tag{4.7}$$

an.

§5. Um endlich (4.7) in die Gestalt (1.2) zu bringen, spalten wir das erste Glied von $S^{\mu\nu}$, nämlich $T^{(\mu|\varepsilon|\nu)}{}_{;\varepsilon}$, von den anderen beiden ab. Da offenbar, infolge von (1.4),

$$F^{\sigma_1\cdots\sigma_{j+1}}{}_{;\sigma_{j+1}} = - F^{\sigma_1\cdots\sigma_j} + E^{\sigma_1\cdots\sigma_j} \tag{5.1}$$

ist, haben wir

$$\begin{aligned} T^{\mu\varepsilon\nu}{}_{;\varepsilon} = \frac{1}{2}\sum_{j=0}^{s-1}\sum_{i=1}^{j+4}\Big[&(E^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_{j+4}} - F^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_{j+4}})B_{\sigma_1\cdots;\cdots\sigma_{i-1}\cdot}{}^{\nu}{}_{\sigma_{i+1}\cdots\sigma_{j+4}} \\ &+ F^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_{j+5}}B_{\sigma_1\cdots;\cdots\sigma_{i-1}\cdot}{}^{\nu}{}_{\sigma_{i+1}\cdots\sigma_{j+5}}\Big] \\ = \frac{1}{2}\sum_{j=0}^{s}\sum_{i=1}^{j+4} & E^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_{j+4}}B_{\sigma_1\cdots;\cdots\sigma_{i-1}\cdot}{}^{\nu}{}_{\sigma_{i+1}\cdots\sigma_{j+4}} \\ - \frac{1}{2}\sum_{i=1}^{4} & F^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_4}B_{\sigma_1\cdots;\cdots\sigma_{i-1}\cdot}{}^{\nu}{}_{\sigma_{i+1}\cdots\sigma_4} - \frac{1}{2}\sum_{j=0}^{s-1}F^{\sigma_1\cdots\sigma_{j+4}\mu}B_{\sigma_1\cdots;\cdots\sigma_{j+4}\cdot}{}^{\nu} . \end{aligned} \tag{5.2}$$

Unter Anwendung der Symmetrieeigenschaften von $B_{\mu\nu\sigma\rho}$ lässt sich unschwer nachweisen, dass sich die zweite Summe in (5.2) einfach als $\frac{4}{3}B^{\nu}_{\cdot\alpha\sigma\varrho}X^{\mu\alpha\sigma\varrho}$ schreiben lässt. Berücksichtigt man noch (2.9) und (1.7), so erhält man schliesslich aus (5.2)

$$T^{\mu\varrho\nu}{}_{;\varrho} = -\eta^{\mu\nu} - \tfrac{2}{3}B^{\nu}_{\cdot\alpha\sigma\varrho}X^{\mu\alpha\sigma\varrho} + t^{\mu\nu} . \qquad (5.4)$$

Setzt man diese Beziehung in (4.7) ein, so ergibt sich (1.2), was zu beweisen war.

§6. Wir wollen jetzt durch direkte Rechnung nachweisen, dass die Divergenz der rechten Seite von (1.2) identisch verschwindet. Zu diesem Zweck beweisen wir zunächst die Gleichungen

$$X^{\mu\nu\sigma\varrho}{}_{;\varrho\sigma\nu} = B^{\mu}_{\cdot\alpha\sigma\varrho}X^{\lambda\alpha\sigma\varrho}{}_{;\lambda} + \tfrac{1}{3}B^{\mu}_{\cdot\alpha\sigma\varrho;\lambda}X^{\lambda\alpha\sigma\varrho} - \tfrac{1}{3}(B^{\lambda}_{\cdot\alpha\sigma\varrho}X^{\mu\alpha\sigma\varrho})_{;\lambda} \quad (6.1)$$

und $$t^{\nu\mu}{}_{;\nu} = -\tfrac{1}{2}K_{;}{}^{\mu} + \tfrac{1}{6}X^{\alpha\beta\sigma\varrho}B_{\alpha\beta\sigma\varrho;}{}^{\mu} - T^{\beta\varepsilon\alpha}B^{\mu}_{\cdot\beta\alpha\varepsilon} . \qquad (6.2)$$

(6.1) kann man beweisen, indem man die drei Ausdrücke $X^{\mu\nu\sigma\varrho}{}_{;\varrho\sigma\nu}$, $X^{\mu\nu\sigma\varrho}{}_{;\varrho\nu\sigma}$, $X^{\mu\nu\sigma\varrho}{}_{;\nu\varrho\sigma}$ addiert, und dabei im Auge behält, dass sich infolge der Identität von Ricci der zweite und dritte Ausdruck, unter Hinzufügen von gewissen Zusatzgliedern, durch den ersten ersetzen lassen. Andererseits kann man aber die Summe der drei Ausdrücke in der Form $(X^{\mu\nu\sigma\varrho} + X^{\mu\sigma\nu\varrho} + X^{\mu\sigma\varrho\nu})_{;\varrho\sigma\nu}$ schreiben, welche wegen (4.4) und (4.41) identisch verschwindet, worauf (6.1) sofort folgt. Was (6.2) betrifft, so haben wir infolge von

(5.1) $$t^{\nu\mu}{}_{;\nu} = -\tfrac{1}{2}\sum_{j=0}^{s-1}\Big[(E^{\sigma_1\cdots\sigma_{j+\nu}} - F^{\sigma_1\cdots\sigma_{j+\nu}})B_{\sigma_1\cdots;\cdots\sigma_{j+\nu};}{}^{\mu}$$
$$+ F^{\sigma_1\cdots\sigma_{j+s}}B_{\sigma_1\cdots;\cdots\sigma_{j+s};}{}^{\mu} + 2g^{\mu\nu}F^{\sigma_1\cdots\sigma_{j+\nu}\varepsilon}B_{\sigma_1\cdots;\cdots\sigma_{j+\nu}[\nu\varepsilon]}\Big]$$
$$= -\tfrac{1}{2}\sum_{j=0}^{s}E^{\sigma_1\cdots\sigma_{j+\nu}}B_{\sigma_1\cdots;\cdots\sigma_{j+\nu};}{}^{\mu} + \tfrac{1}{2}F^{\sigma_1\cdots\sigma_{\alpha}}B_{\sigma_1\cdots\sigma_{\nu};}{}^{\mu}$$
$$-\tfrac{1}{2}\sum_{j=0}^{s-1}\sum_{i=1}^{j+\alpha}F^{\sigma_1\cdots\sigma_{j+\nu}\varepsilon}B_{\sigma_1\cdots;\cdots\sigma_{i-1}\alpha\sigma_{i+1}\cdots\sigma_{j+\nu}}B_{\sigma_i\cdot\varepsilon\cdot}^{\cdot\mu\cdot\alpha} ,$$

was in der Tat der rechten Seite von (6.2) entspricht.

Bilden wir nun die Divergenz von (1.2) und machen von (6.1) Gebrauch, so erhalten wir

$$P^{\mu\nu}{}_{;\nu} = \tfrac{1}{3}B^{\mu}_{\cdot\alpha\sigma\varrho}X^{\lambda\alpha\sigma\varrho}{}_{;\lambda} - \tfrac{1}{3}(B^{\lambda}_{\cdot\alpha\sigma\varrho}X^{\mu\alpha\sigma\varrho})_{;\lambda} + \tfrac{1}{2}K_{;}{}^{\mu}$$
$$+ (T^{\varepsilon(\mu\nu)} - T^{(\mu\nu)\varepsilon})_{;\varepsilon\nu} + t^{(\mu\nu)}{}_{;\nu} . \qquad (6.3)$$

Nun ist aber

$$(T^{\varepsilon(\mu\nu)} - T^{(\mu\nu)\varepsilon})_{;\varepsilon\nu} = T^{[\varepsilon|\nu|\mu]}{}_{;\nu\varepsilon} + T^{\varepsilon\mu\nu}{}_{;[\varepsilon\nu]} + 2T^{[\varepsilon|\nu|\mu]}{}_{;[\varepsilon\nu]}$$
$$= (\tfrac{2}{3}B^{[\varepsilon}{}_{\cdot\alpha\sigma\varrho}X^{\mu]\alpha\sigma\varrho} + t^{[\varepsilon\mu]})_{;\varepsilon} + T^{\varepsilon\mu\nu}{}_{;[\varepsilon\nu]} + 2T^{[\varepsilon|\nu|\mu]}{}_{;[\varepsilon\nu]},$$

(wegen (5.4) und (2.91)). Also wird (6.3) jetzt zu

$$P^{\mu\nu}{}_{;\nu} = -\tfrac{1}{3}X^{\varepsilon\alpha\sigma\varrho}B^{\mu}{}_{\cdot\alpha\sigma\varrho;\varepsilon} + \tfrac{1}{2}K_{;}{}^{\mu} + t^{\nu\mu}{}_{;\nu} + T^{\varepsilon\mu\nu}{}_{;[\varepsilon\nu]} + 2T^{[\varepsilon|\nu|\mu]}{}_{;[\varepsilon\nu]},$$

was sich wegen (6.2) auf

$$P^{\mu\nu}{}_{;\nu} = \tfrac{1}{6}X^{\varepsilon\alpha\sigma\varrho}(B_{\varepsilon\alpha\sigma\varrho;}{}^{\mu} - 2B_{\sigma\varrho}{}^{\mu}{}_{\alpha;\varepsilon}) + (T^{\varepsilon\mu\nu} + 2T^{[\varepsilon|\nu|\mu]})_{;[\varepsilon\nu]} - T^{\beta\varepsilon\alpha}B^{\mu}{}_{\cdot\beta\alpha\varepsilon}$$

reduziert. Wenden wir die Bianchische Identität auf den ersten und die Ricci'sche auf den zweiten Klammerausdruck an, so ergibt sich $P^{\mu\nu}{}_{;\nu} = \tfrac{1}{3}X^{\varepsilon\lambda\sigma\varrho}B^{\mu}{}_{\cdot\sigma\lambda(\varepsilon;\varrho)} + \tfrac{1}{2}T^{\varepsilon\alpha\nu}(B^{\mu}{}_{\cdot\varepsilon\nu\alpha} + B^{\mu}{}_{\cdot\varepsilon\alpha\nu} + B^{\mu}{}_{\cdot\alpha\nu\varepsilon}) - T^{\beta\varepsilon\alpha}B^{\mu}{}_{\cdot\beta\alpha\varepsilon}$

Das erste Glied verschwindet, da

$$X^{\varepsilon\lambda\sigma\varrho}B^{\mu}{}_{\cdot\sigma\lambda(\varepsilon;\varrho)} = X^{\lambda(\varepsilon\varrho)\sigma}B^{\mu}{}_{\cdot\sigma\lambda\varepsilon;\varrho} = -\tfrac{1}{2}X^{\lambda\sigma\varepsilon\varrho}B^{\mu}{}_{\cdot\sigma\lambda\varepsilon;\varrho} = 0$$

ist, während die übrigen Glieder infolge der Symmetrieeigenschaften von $B_{\mu\nu\sigma\varrho}$ verschwinden. Also ist $P^{\mu\nu}{}_{;\nu} \equiv 0$, was zu beweisen war.

§7. Betrachten wir nun eine relative Fundamentalinvariante $K\gamma$, die dadurch entsteht, dass man in einer im Weylschen Sinne eichinvarianten relativen Fundamentalinvariante die Koeffizienten der linearen Fundamentalform gleich Null setzt, so folgt aus einer bekannten Gleichung (<u>Weyl</u>, 1919, § 34, <u>S</u>. 247, <u>Glg</u>. 74), dass die verjüngte Variationsableitung P $(= P^{\mu}_{\mu})$ von K die Gestalt einer Divergenz haben muss. Behalten wir nun im Auge, dass $B_{\mu\nu\sigma\varrho}$ die entartete Form des entsprechenden Weyl'schen Krümmungstensors ist, so erkennt man leicht, dass sich $K_{(p)}$ mit $\lambda^{-q_{(p)}}$ multipliziert, $(q_{(p)} = \bar{M}_{(p)} - 2M_{(p)})$, wenn wir die $g_{\mu\nu}$ mit einem konstanten Faktor λ^2 versehen. Da sich γ gleichzeitig mit λ^{n} multipliziert, γK aber sich laut Voraussetzung überhaupt nicht ändert, muss die Identität

$$F(\lambda^{-q_{(1)}}K_{(1)},\ldots,\lambda^{-q_{(p)}}K_{(p)},\ \ldots) \equiv \lambda^{-n}F(K_{(1)},\ldots,K_{(p)},\ldots) \qquad (7.1)$$

bestehen, woraus folgt, dass

$$\sum_p q_{(p)}K_{(p)}\frac{\partial K}{\partial K_{(p)}} = nK \qquad (7.2)$$

ist, wenn wir nach λ, (welches wir jetzt als einen variablen Parameter ansehen), differenzieren, und sodann $\lambda = 1$ setzen.

Multiplizieren wir aber (1.2) mit $g_{\mu\nu}$, so ist es in der Tat möglich, P in die Form

$$P = \tfrac{1}{2}\Big(nK - \sum_p q_{(p)}K_{(p)}\frac{\partial K}{\partial K_{(p)}}\Big) + \ldots \qquad (7.3)$$

zu bringen, wo die durch Punkte angedeuteten Glieder schon die Gestalt einer Divergenz haben. Da wir natürlich voraussetzen, dass K selbst keine Divergenz ist, (im anderen Falle würde ja $P^{\mu\nu}$ identisch verschwinden), so sehen wir, dass P nur eine Divergenz sein kann, wenn das in (7.3) ausgeschriebene Glied verschwindet, also gerade die Bedingung (7.2) erfüllt ist.

Umgekehrt erkannen wir also, dass der aus der Variationsableitung einer Fundamentalinvariante K entspringende Skalar P nur dann die Form einer Divergenz besitzt, wenn γK die entartete Form einer im Weylschen Sinne eichinvarianten Dichte ist.

§8.(a) Ist K eine Funktion der $B_{\mu\nu\sigma\varrho}$ und $g_{\mu\nu}$ allein, ist also durchwegs $s = 0$, so bleiben in (1.2) nur die ersten drei Glieder bestehen, wovon jetzt das zweite auch schon ohne die μ und ν einschliessenden Klammern symmetrisch ist. Auch erkennt man unschwer, dass einfach

$$\tfrac{1}{2}X^{\mu\nu\sigma\varrho} \equiv Z^{\mu\nu\sigma\varrho} = \frac{\partial K}{\partial g_{\mu\nu,\sigma\varrho}} \qquad (8.1)$$

ist. Also ist in diesem Falle

$$P^{\mu\nu} = Z^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} - \tfrac{2}{3}B^{\mu}{}_{\cdot\alpha\sigma\varrho}Z^{\nu\alpha\sigma\varrho} + \tfrac{1}{2}g^{\mu\nu}K \quad . \tag{8.2}$$

(b) Schliesslich betrachten wir noch ein Beispiel einer einfachen Invariante höherer Ordnung, und zwar die quadratische Invariante

$$K = G_{;\varepsilon}G^{;\varepsilon} \quad , \ (s=1), \tag{8.3}$$

wobei G die skalare Krümmung $g^{\mu\nu}G_{\mu\nu} = g^{\mu\nu}g^{\sigma\varrho}B_{\mu\nu\sigma\varrho}$ bedeutet. (Wir schreiben das Semikolon der Einfachkeit halber oben an). Wir haben sofort

$$E^{\mu\nu\sigma\varrho\lambda} = F^{\mu\nu\sigma\varrho\lambda} = 2g^{\mu\nu}g^{\sigma\varrho}G^{;\lambda} \quad . \tag{8.4}$$

Setzt man für einen beliebigen Ausdruck A

$$g^{\alpha\beta}A_{;\alpha\beta} = \Box A \quad ,$$

so folgt

$$\begin{aligned} F^{\mu\nu\sigma\varrho} &= -2g^{\mu\nu}g^{\sigma\varrho}\Box G \quad , \\ X^{\mu\nu\sigma\varrho} &= -8\,g^{[\mu[\nu}g^{\sigma]\varrho]}\Box G \quad , \end{aligned} \tag{8.5}$$

und

$$\begin{aligned} T^{\alpha\beta\gamma} &= 4G^{\alpha\gamma;\beta} , \\ t^{\mu\nu} &= -\,G^{;\mu}G^{;\nu} \quad . \end{aligned} \tag{8.6}$$

Einsetzen in (1.2) ergibt jetzt sofort

$$\tfrac{1}{2}P^{\mu\nu} = (\Box G)^{;\mu\nu} + G^{\mu\nu}\Box G - \tfrac{1}{2}G^{;\mu}G^{;\nu} - g^{\mu\nu}(\Box^2 G - \tfrac{1}{4}G_{;\varepsilon}G^{;\varepsilon}) \ , \tag{8.7}$$

($\Box^2 G$ bedeutet hierbei $\Box(\Box G)$). Es ist zu beachten, dass auch die Variationsableitung der Invariante $-G\Box G$, (s=2), durch (8.7) dargestellt wird, wie man durch partielle Integration, auf das Integral $-\int \gamma G \Box G d\tau$ angewendet, leicht beweist.

(c) Für den Fall der eben betrachteten Invariante $G_{;\varepsilon}G^{;\varepsilon}$ ist $\overline{M} = 10$ und $M = 2$, also $q = 6$, sodass laut (7.2) P eine Divergenz sein muss, wenn $n = 6$ ist. In der Tat ist dann

$$P = \Box(G^2 - 10\,\Box G) \quad . \tag{8.8}$$

Bibliographie.

Weyl, H. 1920 Raum, Zeit, Materie. Dritte Auflage. Julius Springer. Berlin

Schouten, J.A. 1924 Der Ricci-Kalkül. Julius Springer. Berlin.

Veblen, O. 1927 Invariants of Quadratic Differential Forms. Cambridge Tracts. No. 24.

Eddington, A. 1930 The Mathematical Theory of Relativity. Zweite Auflage. Cambridge University Press.

ON THE HAMILTONIAN DERIVATIVES ARISING FROM A CLASS OF GAUGE-INVARIANT ACTION PRINCIPLES IN A W_n.

by H. A. Buchdahl, Department of Physics, University of Tasmania.

§1. The setting up of field equations in Weyl's unified field theory [Weyl(1), Eddington(1)] on the basis of a postulated action principle requires the determination of Hamiltonian derivatives with respect to the functions $g_{\mu\nu}$ and k_μ, where the $g_{\mu\nu}$ are the components of the metrical tensor, and the k_μ are the coefficients of the fundamental linear form. In a W_n, i.e. an n-dimensional space whose coefficients of affine connection are given by

$$\Gamma^\lambda_{\mu\nu} = \{\mu\nu,\lambda\} - (\delta^\lambda_\mu k_\nu + \delta^\lambda_\nu k_\mu - g_{\mu\nu}k^\lambda), \qquad (1.1)$$

($\{\mu\nu,\lambda\}$ are the usual Christoffel symbols formed with respect to the $g_{\mu\nu}$; and $k^\lambda = g^{\lambda\alpha}k_\alpha$), let K be a gauge-covariant[1] fundamental invariant of gauge-weight $-n/2$, so that $K\sqrt{-g}$, ($g = |g_{\mu\nu}|$), is a gauge-invariant tensor density of coordinate-weight +1. Then the Hamiltonian derivatives of K with respect to $g_{\mu\nu}$ and k_μ, $P^{\mu\nu}$ and Q^μ say, are defined by the equation

---

1. We say that a tensor or tensor-density is gauge-covariant of gauge-weight w if in a gauge transformation it becomes multiplied by a factor λ^{2w} when the $g_{\mu\nu}$ are replaced by $\lambda^2 g_{\mu\nu}$ and the k_μ by $k_\mu + (\log\lambda)_{,\mu}$, λ being an arbitrary function of the coordinates. When $w=0$ we speak of gauge-invariance. Also we replace the more usual term 'weight' as applied to a tensor-density by the term 'coordinate-weight' to avoid confusion with the gauge-weight.

$$\delta\int K\sqrt{-g}\,d\tau \equiv \delta J = \int (P^{\mu\nu}\delta g_{\mu\nu} + Q^{\mu}\delta k_{\mu})\sqrt{-g}\,d\tau \quad , \qquad (1.2)$$

($d\tau = dx_1 \ldots dx_n$), all variations of the $g_{\mu\nu}$, k_{μ} and their derivatives up to the required order vanishing on the boundary of the region of integration. Henceforth we restrict ourselves to the case in which K contains the derivatives of the $g_{\mu\nu}$ only up to the second order, (and those of the k_{μ} only up to the first order).

The main object of this paper is to express $P^{\mu\nu}$ and Q^{μ} in terms of the derivatives of K with respect to the highest derivatives of $g_{\mu\nu}$ and k_{μ} occurring in K, i.e. in terms of the gauge-covariant tensors

$$Z^{\mu\nu\sigma\rho} = \frac{\partial K}{\partial g_{\mu\nu,\sigma\rho}} \qquad (1.31)$$

and

$$S^{\mu\nu} = \frac{\partial K}{\partial k_{\mu,\nu}} \quad . \qquad (1.32)$$

(Subscripts following a comma denote ordinary differentiation). The result to be established is

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3}Z^{\sigma\rho\alpha(\mu}B_{\sigma\rho\alpha\cdot}{}^{\nu)} + \tfrac{1}{2}g^{\mu\nu}K - \tfrac{1}{2}S^{\sigma(\mu}F_{\sigma\cdot}{}^{\nu)} \quad , \qquad (1.4)$$

$$Q^{\mu} = -S^{\mu\nu}{}_{;\nu} \quad . \qquad (1.5)$$

Here we have used curved and square brackets to denote mixing and alternating over indices respectively [Schouten(1)]. $B_{\mu\nu\sigma\cdot}{}^{\varepsilon}$ is the gauge-invariant curvature tensor of Weyl's theory, viz.

$$\tfrac{1}{2}B_{\mu\nu\sigma\cdot}{}^{\varepsilon} = \Gamma^{\varepsilon}_{\mu[\sigma,\nu]} + \Gamma^{\lambda}_{\mu[\sigma}\Gamma^{\varepsilon}_{\nu]\lambda} \quad ; \qquad (1.6)$$

and

$$F_{\mu\nu} = 2k_{[\mu,\nu]} \quad . \qquad (1.7)$$

Subscripts after a semicolon denote gauge-invariant covariant differentiation, defined for a tensor $T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k}$ of gauge-weight w, as

$$T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k;\lambda} = T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k,\lambda} - \sum_{s=1}^{k}\Gamma^{\varepsilon}_{\nu_s\lambda}T^{\mu\cdots\mu_i}_{\nu_1\cdots\nu_{s-1}\varepsilon\nu_{s+1}\cdots\nu_k}$$
$$+ \sum_{s=1}^{i}\Gamma^{\mu_s}_{\varepsilon\lambda}T^{\mu_1\cdots\mu_{s-1}\varepsilon\mu_{s+1}\cdots\mu_i}_{\nu_1\cdots\nu_k} - 2w\,k_{\lambda}T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k} \quad . \qquad (1.8)$$

Note that $g_{\mu\nu;\lambda}$ and $g^{\mu\nu}{}_{;\lambda}$ vanish identically, so that indices may be raised and lowered with the same ease as in Riemannian geometry. (1.4) and (1.5) make the transformation characters of $P^{\mu\nu}$ and Q^{μ} obvious, i.e. $P^{\mu\nu}$ is a symmetrical contravariant tensor of gauge-weight $-(n+2)/2$, whilst the contravariant vector Q^{μ} has the gauge-weight $-n/2$.

When the form of K is known explicitly the evaluation of $P^{\mu\nu}$ and Q^{μ} by means of (1.4), (1.5) is a matter of relative ease. For $Z^{\mu\nu\sigma\varrho}$ and $S^{\mu\nu}$ are connected with the tensor[2].

$$E^{\mu\nu\sigma\varrho} = \frac{\partial K}{\partial B_{\mu\nu\sigma\varrho}} \qquad (1.91)$$

(which may usually be determined by inspection) through the relations

$$\left.\begin{aligned} Z^{\mu\nu\sigma\varrho} &= Y^{\mu\nu(\sigma\varrho)} \\ S^{\mu\nu} &= -2Y_{\alpha}^{\cdot\alpha\nu\mu} \quad , \end{aligned}\right\} \qquad (1.92)$$

where $$Y^{\mu\nu\sigma\varrho} = E^{[\mu[\nu\sigma]\varrho]} + E^{[\nu[\mu\sigma]\varrho]} + E^{(\mu[\sigma\varrho]\nu)} \quad . \qquad (1.93)$$

Certain identities involving $P^{\mu\nu}$ and Q^{μ} are dealt with in §6, and some explicit examples of Hamiltonian derivatives are considered in §7.

§2. (a) Every fundamental invariant K of the kind under consideration is a function of a number of invariants $K_{(1)}, K_{(2)}, \ldots$,

$$K = F(K_{(1)}, K_{(2)}, \ldots) \quad , \qquad (2.1)$$

each of which may formally be written as a product of components of the contravariant metrical tensor $g^{\mu\nu}$ and of the covariant curvature tensor $B_{\mu\nu\sigma\varrho}$; (see also §5(a)). We write for the p-th of these invariants $$K_{(p)} = \Gamma_{(p)} C_{(p)} \quad , \qquad (2.2)$$

---

2. Symmetry properties of $B_{\mu\nu\sigma\varrho}$ are to be disregarded in this differentiation.

where we have suppressed the indices of the tensors $\Gamma_{(p)}$ and $C_{(p)}$. $C_{(p)}$ is the product of m_p factors $B_{\cdots\cdot}$, whilst $\Gamma_{(p)}$ is the product of $2m_p$ factors $g^{\cdot\cdot}$, where m_p is a positive integer. $\Gamma_{(p)}$ and $C_{(p)}$ are each of rank $4m_p$, their respective gauge-weights being $-2m_p$ and $+m_p$, so that the gauge-weight of $K_{(p)}$ is $-m_p$. Since K has, by hypothesis, the gauge-weight $-n/2$, (2.1) at once gives rise to the identity

$$F(\lambda^{-2m_1}K_{(1)}, \lambda^{-2m_2}K_{(2)}, \ldots) \equiv \lambda^{-n}F(K_{(1)}, K_{(2)}, \ldots), \tag{2.31}$$

and therefore
$$\Sigma\, m_p K_{(p)}\frac{\partial K}{\partial K_{(p)}} = nK/2 . \tag{2.32}$$

Now let $E^{\mu\nu\sigma\rho}_{(p)}$ be the sum of the m_p tensors obtained by omitting in turn each single factor from $C_{(p)}$ and in each case replacing the free indices which result by μ, ν, σ, ρ, in their proper order; then[3]
$$E^{\mu\nu\sigma\rho} = \Sigma\, E^{\mu\nu\sigma\rho}_{(p)}\frac{\partial K}{\partial K_{(p)}} = \frac{\partial K}{\partial B_{\mu\nu\sigma\rho}} . \tag{2.4}$$

(b) Suppose $K_{(p)}$ to be written simply as a product of m_p components of the curvature tensor, so that to each superscript there corresponds just one subscript; and denote by $-{}^{\prime}\eta^{\mu\nu}_{(p)}$ the sum of the tensors obtained by taking in turn each superscript and replacing it by μ, at the same time raising the subscript with which it is paired and replacing it by ν. Then in forming the sum

$$\sum_{i=1}^{4} E_{(p)}^{\sigma_1\cdots\sigma_{i-1}\mu\sigma_{i+1}\cdots\sigma_4} B_{\sigma_1\cdots\sigma_{i-1}}{}^{\nu}{}_{\sigma_{i+1}\cdots\sigma_4} \tag{2.51}$$

we in effect go twice through the process just described, except that where a superscript was replaced by μ the first time it is replaced by ν the second time, so that (2.51) is symmetrical, and

---

3. See footnote 2.

equal to $-2'\eta^{\mu\nu}_{(p)}$. Remembering that

$$\delta g^{\mu\nu} = - g^{\mu\sigma} g^{\nu\varrho} \delta g_{\sigma\varrho} \quad , \qquad (2.52)$$

it then follows that

$$\eta^{\mu\nu} = \Sigma\ \eta^{\mu\nu}_{(p)} \frac{\partial K}{\partial K_{(p)}} = \Sigma\ C_{(p)} \frac{\partial \Gamma_{(p)}}{\partial g_{\mu\nu}} \frac{\partial K}{\partial K_{(p)}} = -\frac{1}{2} \sum_{i=1}^{4} E^{\sigma_1 \cdots \sigma_{i-1} \mu \sigma_{i+1} \cdots \sigma_4} B_{\sigma_1 \cdots \sigma_{i-1} \cdot \sigma_{i+1} \cdots \sigma_4}^{\qquad\quad \nu} . \qquad (2.6)$$

We emphasise again that $\eta^{\mu\nu}$, as given by (2.6) is symmetrical as it stands,

$$\eta^{[\mu\nu]} = 0 \quad . \qquad (2.61)$$

Writing $\sqrt{-g} = \gamma$, we now have from (1.2)

$$\begin{aligned} \delta J &= \int (\tfrac{1}{2} g^{\mu\nu} K \delta g_{\mu\nu} + \delta K) \gamma\, d\tau \\ &= \int [(\eta^{\mu\nu} + \tfrac{1}{2} g^{\mu\nu} K) \delta g_{\mu\nu} + E^{\mu\nu\sigma\varrho} \delta B_{\mu\nu\sigma\varrho}] \gamma\, d\tau \quad . \qquad (2.7) \end{aligned}$$

§3. (a) It is convenient to rewrite (2.7) in the form

$$\delta J = \int [(\eta^{\mu\nu} + \tfrac{1}{2} g^{\mu\nu} K + E^{\alpha\beta\sigma\mu} B_{\alpha\beta\sigma\cdot}^{\quad\ \nu}) \delta g_{\mu\nu} + E^{\mu\nu\sigma\cdot}_{\quad\ \varepsilon} \delta B_{\mu\nu\sigma\cdot}^{\quad\ \varepsilon}] \gamma\, d\tau \quad . \qquad (3.1)$$

By (1.6),

$$\begin{aligned} \tfrac{1}{2} E^{\mu\nu\sigma\cdot}_{\quad\ \varepsilon} \delta B_{\mu\nu\sigma\cdot}^{\quad\ \varepsilon} &= E^{\mu\nu\sigma\cdot}_{\quad\ \varepsilon} (\delta\Gamma^{\varepsilon}_{\mu[\sigma,\nu]} + \Gamma^{\lambda}_{\mu[\sigma} \delta\Gamma^{\varepsilon}_{\nu]\lambda} + \Gamma^{\varepsilon}_{\lambda[\nu} \delta\Gamma^{\lambda}_{\sigma]\mu}) \\ &= E^{\mu[\sigma\nu]\cdot}_{\qquad \varepsilon} \delta\Gamma^{\varepsilon}_{\mu\nu,\sigma} + (\Gamma^{\mu}_{\lambda\sigma} E^{\lambda[\nu\sigma]\cdot}_{\qquad \varepsilon} + \Gamma^{\lambda}_{\varepsilon\sigma} E^{\mu[\sigma\nu]\cdot}_{\qquad \lambda}) \delta\Gamma^{\varepsilon}_{\mu\nu} . \end{aligned}$$

Hence, on integrating by parts,

$$\begin{aligned} \delta J' &\equiv \tfrac{1}{2} \int E^{\mu\nu\sigma\cdot}_{\quad\ \varepsilon} \delta B_{\mu\nu\sigma\cdot}^{\quad\ \varepsilon} \gamma\, d\tau \\ &= \int [\gamma^{-1} (\gamma E^{\mu[\nu\sigma]\cdot}_{\qquad \varepsilon})_{,\sigma} + \Gamma^{\mu}_{\lambda\sigma} E^{\lambda[\nu\sigma]\cdot}_{\qquad \varepsilon} + \Gamma^{\lambda}_{\varepsilon\sigma} E^{\mu[\sigma\nu]\cdot}_{\qquad \lambda}] \gamma\, \delta\Gamma^{\varepsilon}_{\mu\nu}\, d\tau \quad , \qquad (3.2) \end{aligned}$$

since the integrated part vanishes, by hypothesis. Now the integrand of (3.2) is a gauge-invariant scalar-density; whilst the $\gamma\,\delta\Gamma^{\varepsilon}_{\mu\nu}$ form the components of an arbitrary gauge-invariant tensor-density. The coefficient of $\gamma\,\delta\Gamma^{\varepsilon}_{\mu\nu}$ must therefore be a tensor, from which it follows without further calculation that

$$\delta J' = \int \gamma E^{\mu[\nu\sigma]\cdot}_{\qquad \varepsilon;\sigma} \delta\Gamma^{\varepsilon}_{\mu\nu}\, d\tau \quad , \qquad (3.3)$$

since a number of terms all of which have $\Gamma^{\varepsilon}_{\mu\nu}$ as factors cannot alone form the components of a tensor.

(b) By (1.1), (2.52) we have

$$\delta J' = \int \gamma E^{\mu[\nu\sigma]\cdot}{}_{\varepsilon;\sigma}\Big[\tfrac{1}{2}g^{\varepsilon\lambda}(\delta g_{\mu\lambda,\nu} + \delta g_{\nu\lambda,\mu} - \delta g_{\mu\nu,\lambda})$$
$$- (\delta^{\varepsilon}_{\mu}\delta k_{\nu} + \delta^{\varepsilon}_{\nu}\delta k_{\mu} - g_{\mu\nu}g^{\varepsilon\lambda}\delta k_{\lambda}) + (k^{\varepsilon}\delta g_{\mu\nu} - k^{\beta}g_{\mu\nu}g^{\alpha\varepsilon}\delta g_{\alpha\beta}) + \ldots\Big]d\tau, \qquad (3.4)$$

where the dots indicate terms having Christoffel three-indices symbols as factors. Changing dummy indices and integrating by parts, we obtain according to the reasoning of §3(a)

$$\delta J' = \int \gamma \Big[\tfrac{1}{2}(E^{\mu[\sigma\lambda]\nu} + E^{\lambda[\sigma\nu]\mu} + E^{\mu[\nu\sigma]\lambda})_{;\sigma\lambda}\,\delta g_{\mu\nu}$$
$$+ (E^{\mu[\sigma\lambda]\cdot}{}_{\lambda} + E^{\cdot\,[\sigma\mu]\lambda}_{\lambda} + E^{\cdot\,[\lambda\sigma]\mu}_{\lambda})_{;\sigma}\,\delta k_{\mu}\Big]d\tau \quad . \qquad (3.5)$$

This may now be substituted in (3.1). By (1.2), (1.93), and remembering that $\delta g_{[\mu\nu]} \equiv 0$, we obtain for $P^{\mu\nu}$ and Q^{μ} the following expressions

$$P^{\mu\nu} = Y^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} + \tfrac{1}{2}g^{\mu\nu}K + t^{\mu\nu} \quad , \qquad (3.6)$$

$$Q^{\mu} = 2Y_{\alpha}^{\cdot\,\alpha\sigma\mu}{}_{;\sigma} \quad , \qquad (3.7)$$

where $t^{\mu\nu} = \frac{1}{2}\sum_{i=1}^{4} e_i E^{\sigma_1\cdots\sigma_{i-1}(\mu|\sigma_{i+1}\cdots\sigma_4|}B_{\sigma_1\ldots\sigma_{i-1}}{}^{\nu)}{}_{\sigma_{i+1}\cdots\sigma_4}$; (3.8)

with $e_i = (-1, -1, -1, +1)$.

§4. (a) In order to bring (3.6), (3.7) into the forms (1.4), (1.5) we note to begin with that

$$Y^{\mu\nu(\sigma\varrho)} = Z^{\mu\nu\sigma\varrho} = \frac{\partial K}{\partial g_{\mu\nu,\sigma\varrho}} \quad . \qquad (4.1)$$

This may be verified by remembering that we must have

$$Z^{\mu\nu[\sigma\varrho]} = Z^{[\mu\nu]\sigma\varrho} = 0 \quad , \qquad (4.21)$$

and that $B_{\mu\nu\sigma\varrho} = 2g_{[\mu[\nu,\sigma]\varrho]} + \cdots$,

where the dots stand for terms not involving second derivatives. From (1.92), (1.93) we may also verify the identities

$$Z^{\mu\nu\sigma\varrho} = Z^{\sigma\varrho\mu\nu} \qquad (4.22)$$

and $Z^{\mu(\nu\sigma\varrho)} = 0$. (4.23)

Considering the first term in (3.6), we have

$$Y^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} = (Y^{\mu\nu(\sigma\varrho)} + Y^{\mu\nu[\sigma\varrho]})_{;\sigma\varrho} = Z^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} + Y^{\mu\nu\sigma\varrho}{}_{;[\sigma\varrho]} \, . \qquad (4.31)$$

The second term of (4.31) may be re-expressed by means of the identity of Ricci, which, by (1.6-8) may be written in the form

$$2T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k;[\sigma\varrho]} = \sum_{s=1}^{i} B_{\varepsilon\varrho\sigma\cdot}^{\mu_s} T^{\mu_1\cdots\mu_{s-1}\varepsilon\mu_{s+1}\cdots\mu_i}_{\nu_1\cdots\nu_k} + \sum_{s=1}^{k} B_{\nu_s\sigma\varrho\cdot}^{\varepsilon} T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_{s-1}\varepsilon\nu_{s+1}\cdots\nu_k} - 2wF_{\sigma\varrho}T^{\mu_1\cdots\mu_i}_{\nu_1\cdots\nu_k} . \quad (4.32)$$

Here $w = -(n+2)/2$, so that

$$Y^{\mu\nu\sigma\varrho}{}_{;[\sigma\varrho]} = B_{\varepsilon\varrho\sigma\cdot}^{(\mu}Y^{\nu)\varepsilon\sigma\varrho} + G_{[\sigma\varrho]}Y^{\mu\nu\sigma\varrho} + \tfrac{1}{2}(n+2)F_{\sigma\varrho}Y^{\mu\nu\sigma\varrho} , \quad (4.33)$$

since $\quad B_{\mu\nu\varepsilon\cdot}^{\varepsilon} = G_{\mu\nu}$, (4.34)

and $\quad B_{\mu(\nu\sigma)\cdot}^{\varepsilon} = 0$. (4.35)

But $\quad G_{[\sigma\varrho]} = -\tfrac{1}{2}nF_{\sigma\varrho}$, (4.36)

and (4.31) reduces to

$$Y^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} = Z^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} + B_{\varepsilon\varrho\sigma\cdot}^{(\mu}Y^{\nu)\varepsilon\sigma\varrho} + F_{\sigma\varrho}Y^{\mu\nu\sigma\varrho} . \quad (4.4)$$

(b) Consider the expression $Y^{\mu\nu\sigma\varrho}B_{\varrho\sigma\nu\cdot}^{\lambda}$. Writing $Y^{\mu\nu\sigma\varrho}$ in terms of $E^{\alpha\beta\lambda\omega}$ by means of (1.93) we may collect into four groups the terms in which the superscript μ occurs in the first, second, third, and fourth places respectively. By means of (4.35) and the cyclic identity $\quad B_{[\mu\nu\sigma]\cdot}^{\varepsilon} = 0$, (4.51)

we find, after some changes in dummy indices,

$$Y^{\mu\nu\sigma\varrho}B_{\varrho\sigma\nu\cdot}^{\lambda} = \tfrac{1}{4}(E^{\mu\nu\sigma\varrho}B_{\varrho\sigma\nu\cdot}^{\lambda} + E^{\nu\mu\sigma\varrho}B_{\sigma\varrho\nu\cdot}^{\lambda} + E^{\nu\sigma\mu\varrho}B_{\sigma\nu\varrho\cdot}^{\lambda} + 3E^{\nu\sigma\varrho\mu}B_{\nu\sigma\varrho\cdot}^{\lambda}). \quad (4.52)$$

Dealing with $Y^{\mu\nu\varrho\sigma}B_{\varrho\sigma\nu\cdot}^{\lambda}$ in an exactly similar manner, we get

$$Y^{\mu\nu\varrho\sigma}B_{\varrho\sigma\nu\cdot}^{\lambda} = \tfrac{1}{2}(E^{\mu\nu\sigma\varrho}B_{\varrho\sigma\nu\cdot}^{\lambda} + E^{\nu\mu\sigma\varrho}B_{\sigma\varrho\nu\cdot}^{\lambda} + E^{\nu\sigma\mu\varrho}B_{\sigma\nu\varrho\cdot}^{\lambda}) . \quad (4.53)$$

From the last two equations the relation

$$E^{\alpha\sigma\varrho(\mu}B_{\alpha\sigma\varrho\cdot}^{\nu)} = \tfrac{2}{3}B_{\varrho\sigma\alpha\cdot}^{(\nu}Z^{\mu)\alpha\sigma\varrho} + 2B_{\varrho\sigma\alpha\cdot}^{(\nu}Y^{\mu)\alpha[\sigma\varrho]} \quad (4.54)$$

follows at once.

(c) The following relations will frequently be needed below, [cf. Eddington(2)]: $\quad B_{\langle\mu|\nu\sigma|\varrho\rangle} = g_{\mu\varrho}F_{\nu\sigma}$, (4.61)

$$B_{\mu\nu\sigma\varrho} - B_{\nu\mu\varrho\sigma} = 4g_{[\mu[\nu}F_{\sigma]\varrho]} + g_{\varrho\mu}F_{\nu\sigma} + g_{\nu\sigma}F_{\varrho\mu} . \quad (4.62)$$

(These correspond to two of the four identities obeyed by the Riemann curvature tensor). Thus we may show with their help that

$$E^{\mu\alpha\sigma\rho}B_{\rho\sigma\alpha\cdot}^{\ \ \ \ \nu} + E^{\alpha\mu\sigma\rho}B_{\sigma\rho\alpha\cdot}^{\ \ \ \ \nu} + E^{\alpha\sigma\mu\rho}B_{\sigma\alpha\rho\cdot}^{\ \ \ \ \nu}$$

$$= E^{\mu\alpha\sigma\rho}B^{\nu}{}_{\cdot\alpha\sigma\rho} + E^{\alpha\mu\sigma\rho}B_{\alpha\cdot\sigma\rho}^{\ \ \nu} + E^{\alpha\sigma\mu\rho}B_{\alpha\sigma\cdot\rho}^{\ \ \ \nu}$$

$$+ F_{\sigma\rho}(E^{\sigma\rho\mu\nu} + E^{\sigma\nu\mu\rho} + E^{\nu\sigma\mu\rho} + E^{\rho\mu\sigma\nu} + E^{\nu\mu\rho\sigma} + E^{\rho\mu\nu\sigma} + 2E^{\mu\rho\sigma\nu})$$

$$+ g_{\sigma\rho}F_{\alpha\cdot}^{\ \nu}(E^{\alpha\mu\sigma\rho} - E^{\rho\mu\sigma\alpha} + E^{\sigma\mu\alpha\rho} + E^{\rho\sigma\mu\alpha} - E^{\sigma\alpha\mu\rho} - E^{\alpha\sigma\mu\rho}). \quad (4.63)$$

This follows simply by applying (4.61), (4.62) and (4.35) to each of the terms on the left hand side, and changing dummy indices. Now add the term $E^{\alpha\sigma\rho\mu}B_{\alpha\sigma\rho\cdot}^{\ \ \ \ \nu}$ to both sides of (4.63) and denote the left hand side of the resulting equation by $s^{\mu\nu}$. Also denote the second set of terms on the right hand side, i.e. $F_{\sigma\rho}(\ldots)$, by $r^{\mu\nu}$. Keeping (1.92), (1.93) and (2.6) in mind, (4.63) now reads

$$s^{\mu\nu} = -2\eta^{\mu\nu} + r^{\mu\nu} - S^{\alpha\mu}F_{\alpha\cdot}^{\ \nu} \quad . \qquad (4.64)$$

Again, by (1.93),

$$r^{(\mu\nu)} = -2F_{\sigma\rho}Y^{\mu\nu\sigma\rho} \quad , \qquad (4.65)$$

whilst the equation

$$s^{\mu\nu} = \tfrac{8}{3}Z^{\mu\alpha\rho\sigma}B_{\sigma\rho\alpha\cdot}^{\ \ \ \ \nu} \qquad (4.66)$$

follows at once by addition of (4.52) and (4.53). Hence, taking the symmetrical part of (4.64), we have, in view of (2.61),

$$\eta^{\mu\nu} = -\tfrac{4}{3}Z^{\sigma\rho\alpha(\mu}B_{\sigma\rho\alpha\cdot}^{\ \ \ \ \nu)} - \tfrac{1}{2}S^{\sigma(\mu}F_{\sigma\cdot}^{\ \nu)} - F_{\sigma\rho}Y^{\mu\nu\sigma\rho} \quad . \qquad (4.67)$$

(d) By (2.6) and (3.8) we have

$$t^{\mu\nu} = \eta^{\mu\nu} + E^{\alpha\sigma\rho(\mu}B_{\alpha\sigma\rho\cdot}^{\ \ \ \ \nu)} \quad . \qquad (4.71)$$

Substituting from (4.54) and (4.67) on the right hand side, this gives

$$t^{\mu\nu} = -\tfrac{4}{3}Z^{\sigma\rho\alpha(\mu}B_{\sigma\rho\alpha\cdot}^{\ \ \ \ \nu)} + \tfrac{2}{3}B_{\rho\sigma\alpha\cdot}^{\ \ \ \ (\nu}Z^{\mu)\sigma\alpha\rho} + 2B_{\rho\sigma\alpha\cdot}^{\ \ \ \ (\nu}Y^{\mu)\alpha[\sigma\rho]}$$

$$- \tfrac{1}{2}S^{\sigma(\mu}F_{\sigma\cdot}^{\ \nu)} - F_{\sigma\rho}Y^{\mu\nu\sigma\rho} \quad . \qquad (4.72)$$

Substituting in (3.6) from (4.4) and (4.72) we obtain, by (4.21), (4.22)

$$P^{\mu\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3}Z^{\sigma\rho\alpha(\mu}B_{\sigma\rho\alpha\cdot}^{\ \ \ \ \nu)} + \tfrac{1}{2}g^{\mu\nu}K - \tfrac{1}{2}S^{\sigma(\mu}F_{\sigma\cdot}^{\ \nu)}$$

$$+ 2B_{\rho\sigma\alpha\cdot}^{\ \ \ \ (\nu}Y^{\mu)\alpha[\sigma\rho]} + B_{\varepsilon\rho\sigma\cdot}^{\ \ \ \ (\mu}Y^{\nu)\varepsilon\sigma\rho} \quad . \qquad (4.73)$$

But the last two terms, may, by (4.35), be written as

$$3B_{[\rho\sigma\alpha]}{}^{(\nu}Y^{\mu)\alpha\sigma\rho} \quad , \tag{4.74}$$

and this vanishes in virtue of (4.51). Hence $P^{\mu\nu}$ takes the form (1.4); which was to be proved.

(e) It remains to show, (in view of (1.32) and (3.7)), that

$$S^{\mu\nu} = \frac{\partial K}{\partial k_{\mu,\nu}} \quad . \tag{4.81}$$

Now [cf. Eddington(2)] $\quad B_{\mu\nu\sigma\rho} = 4g_{[\nu[\rho}k_{\mu],\sigma]} + \ldots \quad ,$ (4.82)

where the dots indicate terms not involving the $k_{\mu,\nu}$. Hence for variations of the $k_{\mu,\nu}$ alone we have

$$\begin{aligned} \delta K &= 4E^{\mu\nu\sigma\rho}\, g_{[\nu[\rho}\delta k_{\mu],\sigma]} = -4E^{[\mu[\nu\sigma]\rho]}g_{\sigma\rho}\,\delta k_{\mu,\nu} \\ &= -2Y^{\sigma\rho\nu\mu}g_{\sigma\rho}\,\delta k_{\mu,\nu} \quad , \end{aligned} \tag{4.83}$$

which is equivalent to (4.81). This completes the proof of the validity of (1.4) and (1.5).

§5. It may happen that a given action principle [§1] has as integrand a (gauge-invariant) scalar-density $\underline{K}$ constructed with the help of the numerical tensor-densities $\varepsilon^{\mu_1\cdots\mu_n}$ and $\varepsilon_{\mu_1\cdots\mu_n}$ of Levi-Civitȧ. Using certain identities involving them [Veblen(1)], it is always possible to bring $\underline{K}$ into the form $K\Upsilon$, where K is of the form considered above. But there is at least one case where it is far simpler to proceed otherwise, viz. when $\underline{K}$ is made up of the $\varepsilon_{\mu_1\cdots\mu_n}$, $\varepsilon^{\mu_1\cdots\mu_n}$ and $B_{\mu\nu\sigma}{}^{\rho}$ alone. We now take the Hamiltonian derivatives of $\underline{K}$ to be tensor-densities defined by

$$\delta\bar{J} = \delta\int \underline{K}d\tau = \int(\underline{P}^{\mu\nu}\delta g_{\mu\nu} + \underline{Q}^{\mu}\delta k_{\mu})d\tau \quad , \tag{5.1}$$

where $\underline{K}$ is of the form

$$\underline{K} = F(\underline{K}_{(1)}, \underline{K}_{(2)}, \ldots) \quad . \tag{5.2}$$

The $\underline{p}$-th of the scalar-densities making up $\underline{K}$, $\underline{K}_{(p)}$ say, is, in general, a product of the $\varepsilon^{\mu_1 \cdots \mu_n}$, $\varepsilon_{\mu_1 \cdots \mu_n}$ and the $B_{\mu\nu\sigma}{}^{\cdot}_{\varepsilon}$. Let there be $\underline{m}_p$ of the latter, some or all of which may appear in one or other of their contacted forms, say $\underline{m}_p - \underline{m}'_p$ of them. Then the $B_{\mu\nu\sigma}{}^{\cdot}_{\varepsilon}$ give rise to $\underline{3m}'_p + \underline{2}(\underline{m}_p - \underline{m}'_p) = \underline{2m}_p + \underline{m}'_p$ free covariant indices, and to $\underline{m}'_p$ free contravariant indices. Hence there must be $(\underline{2m}_p + \underline{m}'_p)/\underline{n}$ of the $\varepsilon^{\mu_1 \cdots \mu_n}$ and $\underline{m}'_p/\underline{n}$ of the $\varepsilon_{\mu_1 \cdots \mu_n}$. It follows that the coordinate weight $\underline{v}_p$ of $\underline{K}_{(p)}$ is $(\underline{2m}_p + \underline{m}'_p - \underline{m}'_p)/\underline{n}$, or

$$v_p = 2m_p/n \quad , \tag{5.21}$$

whilst the gauge-weight is, of course, zero. Considering therefore now a transformation of coordinates (corresponding to the gauge-transformation of §2(a)), we must have

$$F(j^{v_1}\underline{K}_{(1)},\ j^{v_2}\underline{K}_{(2)},\ \ldots) \equiv jF(\underline{K}_{(1)},\ \underline{K}_{(2)},\ \ldots) \quad , \tag{5.31}$$

where $\underline{j}$ is the Jacobian of the transformation; whence, in virtue of (5.21),

$$\Sigma\ m_p \underline{K}_{(p)} \frac{\partial \underline{K}}{\partial \underline{K}_{(p)}} = n\underline{K}/2 \quad . \tag{5.32}$$

(b) We now define a tensor-density $\underline{E}^{\mu\nu\sigma\cdot}{}_{\varrho}$ by means of [4].

$$\underline{E}^{\mu\nu\sigma\cdot}{}_{\varrho} = \frac{\partial \underline{K}}{\partial B_{\mu\nu\sigma}{}^{\cdot}_{\varrho}} \quad . \tag{5.41}$$

(This corresponds exactly to the earlier $E^{\mu\nu\sigma\varrho}$ and the explicit law of formation can be taken over almost unchanged from §2(a)).

Then
$$\delta\bar{J} = \int \underline{E}^{\mu\nu\sigma\cdot}{}_{\varrho}\, \delta B_{\mu\nu\sigma}{}^{\cdot}_{\varrho}\, d\tau \quad . \tag{5.5}$$

The definition of gauge-invariant covariant differentiation may be extended to cover tensor-densities by adding a term $-v\Gamma^{\sigma}_{\sigma\lambda} T^{\mu_1 \cdots \mu_i}_{\nu_1 \cdots \nu_k}$ on the right hand side of (1.8), if $T^{\mu_1 \cdots \mu_i}_{\nu_1 \cdots \nu_k}$ is now a tensor-

---

4. See footnote 2.

density of coordinate-weight $\underline{v}$. Comparing equations (3.1) and (5.5) and considering the steps which led to the result (3.6), (3.7) it is not difficult to see that we now have

$$\left.\begin{aligned} \underline{P}^{\mu\nu} &= \underline{Y}^{\mu\nu\sigma\varrho}{}_{;\sigma\varrho} , \\ \underline{Q}^{\mu} &= 2\underline{Y}_{\alpha}^{\cdot\alpha\sigma\mu}{}_{;\sigma} , \end{aligned}\right\} \quad (5.6)$$

where the $\underline{Y}^{\mu\nu\sigma\varrho}$ are formed of the $\underline{E}^{\mu\nu\sigma\varrho}$ in the same way as the $Y^{\mu\nu\sigma\varrho}$ of the $E^{\mu\nu\sigma\varrho}$.

This result could also have been derived by considering the effect on $\underline{K}$ of an arbitrary infinitesimal transformation of coordinates, which leads to the identity

$$g^{\mu\nu}\underline{K} = -2\underline{t}^{\mu\nu} , \quad (5.7)$$

where the $\underline{t}^{\mu\nu}$ are given by (3.8), but with the $E^{\cdots}$ replaced by the $\underline{E}^{\cdots}$. We may note that (5.6) applies in the particular case in which the action integral $\bar{J}$ is the generalised volume [Eddington(3)], i.e.

$$\underline{K} = (-|G_{\mu\nu}|)^{1/2} . \quad (5.8)$$

§6. (a) It is known that certain identities exist [Weyl(2)] on the one hand between the spur P ($= P^{\mu}_{\mu}$) of $P^{\mu\nu}$ and the divergence of Q^{μ}, and on the other hand between Q^{μ} and the divergence of $P^{\mu\nu}$. These identities are a consequence respectively of the general gauge- and coordinate invariance of the action integrals; and in the present notation these may be written

$$Q^{\nu}{}_{;\nu} = 2P , \quad (6.1)$$

$$P^{\mu\nu}{}_{;\nu} = -\tfrac{1}{2}F^{\mu\nu}Q_{\nu} . \quad (6.2)$$

As a partial check on our results (1.4), (1.5) we shall verify these identities by direct calculation. Multiplying (1.4) throughout by $g_{\mu\nu}$,

$$P = Z_{\alpha}{}^{\cdot\alpha\sigma\rho}{}_{;\sigma\rho} - \tfrac{2}{3}Z^{\sigma\rho\alpha\beta}B_{\sigma\rho\alpha\beta} + \tfrac{1}{2}nK - \tfrac{1}{2}S^{\sigma\rho}F_{\sigma\rho} \,. \qquad (6.3)$$

The first term gives

$$Z_{\alpha}{}^{\cdot\alpha\sigma\rho}{}_{;\sigma\rho} = Y_{\alpha}{}^{\cdot\alpha(\sigma\rho)}{}_{;\sigma\rho} = -\tfrac{1}{2}S^{(\sigma\rho)}{}_{;\sigma\rho} = \tfrac{1}{2}Q^{\nu}{}_{;\nu} - \tfrac{1}{2}S^{\sigma\rho}{}_{;[\sigma\rho]} \,.$$

Multiplying (4.33) by $g_{\mu\nu}$ we have

$$\tfrac{1}{2}S^{\sigma\rho}{}_{;[\sigma\rho]} = B_{\epsilon\rho\sigma\alpha}Y^{\alpha\epsilon\sigma\rho} + \tfrac{1}{2}S^{\sigma\rho}F_{\sigma\rho} \,.$$

Applying (4.61) this yields, since $Y^{[\alpha\epsilon\sigma\rho]} = 0$,

$$S^{\sigma\rho}{}_{;[\sigma\rho]} = 0, \qquad (6.31)$$

and therefore $\qquad Z_{\alpha}{}^{\cdot\alpha\sigma\rho}{}_{;\sigma\rho} = \tfrac{1}{2}Q^{\nu}{}_{;\nu} \,. \qquad (6.32)$

If we now multiply (4.67) by $g_{\mu\nu}$ we get

$$\eta = \eta^{\mu}_{\mu} = -\tfrac{4}{3}Z^{\sigma\rho\alpha\beta}B_{\sigma\rho\alpha\beta} - S^{\sigma\rho}F_{\sigma\rho} \,. \qquad (6.33)$$

On the other hand we have, by (2.6),

$$\eta = -2E^{\alpha\beta\sigma\rho}B_{\alpha\beta\sigma\rho} = -2\,\Sigma\,E^{\mu\nu\sigma\rho}_{(p)}B_{\mu\nu\sigma\rho}\frac{\partial K}{\partial K_{(p)}}$$

$$= \Sigma\, m_p K_{(p)}\frac{\partial K}{\partial K_{(p)}} = nK/2 \quad , \text{ by } (2.32); \qquad (6.34)$$

so that the first identity is verified.

(b) The verification of (6.2) is somewhat more troublesome. In order not to lengthen this paper unduly, we shall therefore give only the main steps of the argument, omitting some of the more detailed manipulations. We have[5], from (6.2),

$$P^{\mu\nu}{}_{;\nu} = Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} - \tfrac{2}{3}(Z^{\sigma\rho\alpha(\mu}B_{\sigma\rho\alpha}{}^{\cdot\nu)})_{;\nu} + \tfrac{1}{2}K^{;\mu} - \tfrac{1}{2}(S^{\sigma(\mu}F_{\sigma}{}^{\cdot\nu)})_{;\nu} \,. \qquad (6.4)$$

The first term, z^{μ} say, may be transformed as follows: Write down the sum $\qquad Z^{\mu\nu\sigma\rho}{}_{;\sigma\rho\nu} + Z^{\mu\nu\sigma\rho}{}_{;\sigma\nu\rho} + Z^{\mu\nu\sigma\rho}{}_{;\nu\sigma\rho} \,; \qquad (6.41)$

then we may bring the subscripts of the last two terms of (6.41) into the order σ,ρ,ν , either (i) by a change of dummy indices, in which case (6.41) is seen to be zero in virtue of (4.21) and

5. We sometimes 'raise' a semicolon merely for ease of reading.

(4.23); or (ii) by applying the Ricci identity once to the second term and twice to the third. In this way we obtain

$$\tfrac{3}{2}Z^{\mu} = B^{\mu}_{\cdot\sigma\rho\varepsilon}\, Z^{\varepsilon\rho\sigma\lambda}{}_{;\lambda} + (B^{[\mu}{}_{\cdot\sigma\rho\varepsilon} Z^{\lambda]\sigma\rho\varepsilon})_{;\lambda} + (Z^{\mu\sigma\rho\lambda} F_{\sigma\rho})_{;\lambda} .$$

Hence, by (4.21), (4.22) and (4.35),

$$Z^{\mu} - \tfrac{2}{3}(Z^{\sigma\rho\alpha(\mu} B_{\sigma\rho\alpha}{}^{\lambda)})_{;\lambda} = -\tfrac{2}{3} B^{\mu}_{\cdot\rho\sigma\alpha;\lambda} Z^{\lambda\rho\sigma\alpha} + \tfrac{1}{2}(S^{\sigma[\mu} F_{\sigma}{}^{\lambda]})_{;\lambda} \quad . \quad (6.42)$$

Now, by (4.54), $K^{;\mu} = E^{\alpha\beta\sigma\rho} B_{\alpha\beta\sigma\rho;}{}^{\mu} = \tfrac{2}{3} Z^{\lambda\nu\sigma\rho} B_{\rho\sigma\nu\lambda;}{}^{\mu} + 2Y^{\lambda\nu[\sigma\rho]} B_{\rho\sigma\nu\lambda;}{}^{\mu}$,

which can be brought into the form

$$K^{;\mu} = \tfrac{2}{3} Z^{\lambda\nu\sigma\rho} B_{\rho\sigma\nu\lambda;}{}^{\mu} + \tfrac{1}{2} S^{\sigma\rho} F_{\sigma\rho;}{}^{\mu} \qquad (6.5)$$

by means of (4.35), (4.51) and (4.61). If we apply the identity of Bianchi, viz.

$$B_{\mu[\nu\sigma\cdot;\lambda]}^{\ \ \ \ \ \varepsilon} = 0 \quad , \qquad (6.6)$$

to the first term and make use of the symmetry properties of $B_{\mu\nu\sigma\rho}$ and $Z^{\mu\nu\sigma\rho}$ we find

$$K^{;\mu} = \tfrac{4}{3} B^{\mu}_{\cdot\rho\sigma\alpha;\lambda} Z^{\lambda\rho\sigma\alpha} - S^{(\sigma\rho)} F^{\mu}_{\cdot\sigma;\rho} + \tfrac{1}{2} S^{\sigma\rho} F_{\sigma\rho;}{}^{\mu} \quad . \qquad (6.7)$$

Substituting from (6.42) and (6.7) in (6.4) we get

$$\begin{aligned} P^{\mu\nu}{}_{;\nu} &= \tfrac{1}{2}(S^{\sigma[\mu} F_{\sigma}{}^{\rho]})_{;\rho} - \tfrac{1}{2} S^{(\sigma\rho)} F^{\mu}_{\cdot\sigma;\rho} + \tfrac{1}{4} S^{\sigma\rho} F_{\sigma\rho;}{}^{\mu} - \tfrac{1}{2}(S^{\sigma(\mu} F_{\sigma}{}^{\rho)})_{;\rho} \\ &= -\tfrac{1}{2} F^{\mu\nu} Q_{\nu} + \tfrac{1}{4} S^{\sigma\rho} g^{\mu\lambda} F_{[\sigma\rho;\lambda]} \quad . \end{aligned} \qquad (6.8)$$

But $\qquad F_{[\sigma\rho;\lambda]} = F_{[\sigma\rho,\lambda]} = 0 \quad , \qquad (6.81)$

by (1.7); and hence (6.2) is proved.

§7. In this paragraph we consider a few examples of explicitly given invariants.

(a) The linear invariant $K = G$, so that we must have $n = 2$.

Then $\qquad E^{\mu\nu\sigma\rho} = g^{\mu\nu} g^{\sigma\rho} \quad , \qquad (7.11)$

whence $\qquad Y^{\mu\nu\sigma\rho} = 2g^{[\mu[\nu} g^{\sigma]\rho]} \quad ,$

$$\left.\begin{aligned} &\text{so that} \qquad Z^{\mu\nu\sigma\rho} = Y^{\mu\nu\sigma\rho} \\ &\text{and} \qquad S^{\mu\nu} = -2g^{\mu\nu} . \end{aligned}\right\} \quad (7.12)$$

(1.4), (1.5) now give

$$\left.\begin{aligned} P^{\mu\nu} &= -G^{(\mu\nu)} + \tfrac{1}{2}g^{\mu\nu}G \ , \\ Q^{\mu} &= 0 \ . \end{aligned}\right\} \quad (7.13)$$

Note that in this case the content of (6.2) also follows directly from (6.6) in the usual way.

(b) The quadratic invariant[6]. $K = G^2 = K_1$, say; $(n = 4)$

Then $$E^{\mu\nu\sigma\rho} = 2g^{\mu\nu}g^{\sigma\rho}G \ , \quad (7.21)$$

whence $$Y^{\mu\nu\sigma\rho} = 4g^{[\mu[\nu}g^{\sigma]\rho]}G \ ,$$

so that $$\left.\begin{aligned} Z^{\mu\nu\sigma\rho} &= Y^{\mu\nu\sigma\rho} \ , \\ \text{and} \quad S^{\mu\nu} &= -12g^{\mu\nu}G \ . \end{aligned}\right\} \quad (7.22)$$

Therefore, by (1.4), (1.5),

$$\left.\begin{aligned} P_1^{\mu\nu} &= -2G^{;(\mu\nu)} - 2GG^{(\mu\nu)} + g^{\mu\nu}(2\,\Box G + \tfrac{1}{2}G^2) \ , \\ Q_1^{\mu} &= 12G^{;\mu} \ , \end{aligned}\right\} \quad (7.23)$$

where the symbol $\Box$ denotes the operation $g^{\alpha\beta}(\ldots)_{;\alpha\beta}$.

(c) Let $K = G_{\mu\nu}G^{\mu\nu} = K_2$, say; $(n = 4)$. (7.3)

Then $$E^{\mu\nu\sigma\rho} = 2g^{\sigma\rho}G^{\mu\nu} \ , \quad (7.31)$$

whence $$Y^{\mu\nu\sigma\rho} = g^{\sigma\rho}G^{(\mu\nu)} + g^{\mu\nu}G^{\rho\sigma} - g^{\mu\sigma}G^{(\rho\nu)} - g^{\nu\sigma}G^{(\rho\mu)} \ ,$$

so that $$\left.\begin{aligned} Z^{\mu\nu\sigma\rho} &= g^{\sigma\rho}G^{(\mu\nu)} + g^{\mu\nu}G^{(\sigma\rho)} - g^{\mu\sigma}G^{(\rho\nu)} - g^{\nu\sigma}G^{(\rho\mu)} \ , \\ \text{and} \quad S^{\mu\nu} &= -4G^{(\mu\nu)} - 2g^{\mu\nu}G + 16F^{\mu\nu} \ . \end{aligned}\right\} \quad (7.32)$$

Substitution in (1.4) and (1.5) gives after some rearranging

$$\left.\begin{aligned} P_2^{\mu\nu} &= \Box G^{(\mu\nu)} - G^{;(\mu\nu)} + 2J^{(\mu;\nu)} - 2B^{(\mu\nu)\sigma\rho}G_{(\sigma\rho)} + 8E^{\mu\nu} + \tfrac{1}{2}g^{\mu\nu}(\Box G + G^{(\sigma\rho)}G_{\sigma\rho}) \\ Q_2^{\mu} &= 4G^{;\mu} - 20J^{\mu} \ , \end{aligned}\right\} \quad (7.33)$$

where we have used the abbreviations

$$J^{\mu} = F^{\mu\nu}{}_{;\nu} \ , \quad (7.34)$$

$$E^{\mu\nu} = -F^{\mu\sigma}F^{\nu}{}_{\cdot\sigma} + \tfrac{1}{4}g^{\mu\nu}F^{\sigma\rho}F_{\sigma\rho} \ . \quad (7.35)$$

---

6. This occurs in Weyl's Action Principle, [Eddington(4)].

(d) Let $K = B_{\mu\nu\sigma\rho} B^{\mu\nu\sigma\rho} = K_3$, say; $(n = 4)$. (7.4)

Then $E^{\mu\nu\sigma\rho} = 2B^{\mu\nu\sigma\rho}$, (7.41)

whence, making use of (4.61), it follows that

$$Y^{\mu\nu\sigma\rho} = 4B^{\sigma\rho(\mu\nu)} + 6g^{\mu\nu}F^{\sigma\rho} + 4g^{\rho(\mu}F^{\nu)\sigma} .$$

Therefore $Z^{\mu\nu\sigma\rho} = 4B^{\sigma\rho(\mu\nu)} + 3g^{(\mu\sigma}F^{\nu)\rho} - 3g^{(\mu\nu}F^{\rho)\sigma}$,

and $S^{\mu\nu} = -8G^{(\mu\nu)} + 24F^{\mu\nu}$. (7.42)

The somewhat tedious rearrangement following substitution of (7.42) in (1.4) is greatly facilitated by observing that

$$\left.\begin{aligned} -2F_{\sigma\rho} B^{\sigma\rho(\mu\nu)} &= F_{\sigma\rho} B^{(\mu|\sigma\rho|\nu)} = g^{\mu\nu}F_{\sigma\rho} F^{\sigma\rho} , \\ F_{\sigma\rho} B^{(\mu\nu)\sigma\rho} &= 2E^{\mu\nu} , \end{aligned}\right\} \quad (7.43)$$

by (4.61) and (4.62). The final result is

$$\left.\begin{aligned} P_3^{\mu\nu} &= 4\Box G^{(\mu\nu)} - 2G^{;(\mu\nu)} - 4B^{(\mu\nu)\sigma\rho}G_{(\sigma\rho)} + 2G^{\mu}{}_{\epsilon}G^{(\nu\epsilon)} + 2G^{\nu}{}_{\epsilon}G^{(\mu\epsilon)} \\ &\quad - 2B_{\sigma\rho\alpha}{}^{\mu}B^{\sigma\rho\alpha\nu} + \tfrac{1}{2}g^{\mu\nu}K_3 + 8J^{(\mu;\nu)} + 8E^{\mu\nu} , \\ Q_3^{\mu} &= 4G^{;\mu} - 32J^{\mu} . \end{aligned}\right\} \quad (7.44)$$

(e) Finally, let $K = F_{\sigma\rho}F^{\sigma\rho} = K_4$, say; $(n=4)$. (7.5)

We may write $K_4 = \frac{1}{16}B_{\alpha\sigma\rho}{}^{\alpha}B_{\beta}{}^{\sigma\rho\beta}$,

from which it follows that

$$E^{\mu\nu\sigma\rho} = \tfrac{1}{8}g^{\mu\rho}g^{\alpha\beta}g^{\nu\lambda}g^{\sigma\epsilon}B_{\alpha\lambda\epsilon\beta} = \tfrac{1}{2}g^{\mu\rho}F^{\nu\sigma} , \quad (7.51)$$

Therefore $Y^{\mu\nu\sigma\rho} = \frac{1}{2}g^{\mu\nu}F^{\sigma\rho}$,

so that $Z^{\mu\nu\sigma\rho} = 0$,

and $S^{\mu\nu} = 4F^{\mu\nu}$, (7.52)

from which it follows at once that

$$\left.\begin{aligned} P_4^{\mu\nu} &= 2E^{\mu\nu} , \\ Q_4^{\mu} &= -4J^{\mu} . \end{aligned}\right\} \quad (7.53)$$

§8. In a W_4 every second order invariant K, admissible in (1.2), which is a rational integral function of the components of the curvature tensor is a linear combination with constant coefficients of the four quadratic invariants considered in the previous paragraph. It is evidemtly possible to eliminate between the $P_i^{\mu\nu}$, $(i=1,2,3,4)$, not only the terms involving third and fourth derivatives of the $g_{\mu\nu}$, but the $E^{\mu\nu}$ and $J^{(\mu;\nu)}$ as well. This result may be put into a convenient form as follows: Let the quadratic second order tensor $L^{\mu\nu}$ be defined by

$$L^{\mu\nu} = GG^{(\mu\nu)} - G^{\mu}_{.\varepsilon}G^{(\nu\varepsilon)} - G^{\nu}_{\varepsilon}G^{(\mu\varepsilon)} - 2B^{(\mu\nu)\sigma\rho}G_{(\sigma\rho)} + B_{\sigma\rho\alpha.}{}^{\mu}B^{\sigma\rho\alpha\nu} - g^{\mu\nu}F_{\sigma\rho}F^{\sigma\rho}, \quad \ldots(8.1)$$

so that $$L = L^{\mu}_{\mu} = K_1 - 4K_2 + K_3 + 12K_4 . \quad (8.2)$$

From §7(b)-(d) we find for the Hamiltonian derivatives $P_L^{\mu\nu}$ and Q_L^{μ} of L

$$-\tfrac{1}{2}P_L^{\mu\nu} = L^{\mu\nu} - \tfrac{1}{4}g^{\mu\nu}L , \quad (8.3)$$

$$Q_L^{\mu} = 0 . \quad (8.4)$$

We show in another paper that $L^{\mu\nu} - \frac{1}{4}g^{\mu\nu}L$ vanishes identically; which constitutes an extension of a result previously obtained by Lanczos, [Lanczos(1)].

References.

Eddington, A. S. (1) The Mathematical Theory of Relativity. 2nd edn. 1930. Cambridge University Prese. §§83 - 90. (2) loc. cit. p. 204, equ. (87.4). (3) loc. cit., p. 232. (4) loc. cit., p. 209ff.

Lanczos, C. (1) Annals of Math. II. 1938. Vol. 39, No. 4. p. 842.

Schouten, J. A. (1) Der Ricci-Kalkül. Julius Springer, Berlin, 1924. pp. 25 and 31.

Veblen, O. (1) Invariants of Quadratic Differential Forms. Cambridge Tracts. No. 24., 1927. p. 5, equs. (3.3), (3.41) and (3.42), and p. 9, equ. (3.4)

Weyl, H. (1) Raum, Zeit, Materie. 3rd edn. Julius Springer, Berlin, 1919. §34ff. (2) loc. cit., pp. 246 - 250.

AN IDENTITY BETWEEN THE HAMILTONIAN DERIVATIVES OF CERTAIN FUNDAMENTAL INVARIANTS IN A W_4.

by H. A. Buchdahl,

Physics Department, University of Tasmania.

§ 1. In a W_4 the most general gauge-invariant fundamental scalar density $\underline{K}$, (not containing the derivatives of the $g_{\mu\nu}$ beyond the second order), which is a rational integral function of the components of the curvature tensor, is

$$\underline{K} = \sqrt{-g}\sum_{i=1}^{4} a_i K_i \quad , \tag{1.1}$$

where

$$\left.\begin{aligned} K_1 &= G^2 \\ K_2 &= G_{\mu\nu}G^{\mu\nu} \\ K_3 &= B_{\mu\nu\sigma\varrho}B^{\mu\nu\sigma\varrho} \\ K_4 &= F_{\mu\nu}F^{\mu\nu} \quad , \end{aligned}\right\} \tag{1.2}$$

and the a_i are constants. The explicit expressions for the Hamiltonian derivatives of these four invariants were obtained by the author in an earlier paper[1] [Buchdahl(1)]. We now show that the Hamiltonian derivatives of the invariant

$$L = K_1 - 4K_2 + K_3 + 12K_4 \tag{1.3}$$

vanish identically. This is an extension of a result previously obtained by Lanczos [Lanczos(1)] for the case of a V_4. It follows that for the purpose of setting up field equations in Weyl's theory on the basis of a gauge-invariant action principle, the Lagrangian

1. This paper will be referred to as H.D. Throughout.

of which is of the form (1.1), only the invariants K_1 , K_2 and K_4 need be considered, that is to say, only invariants which are rational integral functions of the components of the contracted curvature tensor, since we may write[2]

$$K_4 = \tfrac{1}{4} G_{\mu\nu} G^{[\mu\nu]} . \tag{1.4}$$

§ 2. The present proof is made to depend upon the following

Lemma: Let $\varepsilon^{\mu\nu\sigma\varrho}$ be a set of numbers [Veblen(1)] skew-symmetric in every pair of indices, $\varepsilon^{1234} = +1$; and let $A_{\mu\nu\sigma\varrho}$ be a set of quantities possessing the symmetry properties of the Riemann curvature tensor, viz.

$$A_{\mu(\nu\sigma)\varrho} = A_{(\mu|\nu\sigma|\varrho)} = A_{[\mu\nu\sigma]\varrho} = 0 \;;\; A_{\mu\nu\sigma\varrho} = A_{\nu\mu\varrho\sigma} . \tag{2.1}$$

Then, if $\quad t^{\mu}_{\nu} = \varepsilon^{\alpha\beta\eta\zeta}\varepsilon^{\sigma\varrho\lambda\mu} A_{\alpha\beta\sigma\varrho} A_{\eta\zeta\lambda\nu}$, (2.2)

the set of quantities $T^{\mu}_{\nu} = t^{\mu}_{\nu} - \frac{1}{4}\delta^{\mu}_{\nu} t$, $(t = t^{\alpha}_{\alpha})$, (2.3)

vanishes identically.

To prove this consider first the case when $\mu \neq \nu$. Without loss of generality we may take $\mu = 1$, $\nu = 2$, say. Then, constantly keeping (2.1) and the skew-symmetry of $\varepsilon^{\mu\nu\sigma\varrho}$ in mind we have

$$\begin{aligned} t^{1}_{2} = \varepsilon^{\alpha\beta\eta\zeta}\varepsilon^{\sigma\varrho\lambda 1} A_{\alpha\beta\sigma\varrho} A_{\eta\zeta\lambda 2} &= \tfrac{1}{2}\varepsilon^{\alpha\beta\eta\zeta}\varepsilon^{\sigma\varrho\lambda 1} A_{\alpha\beta\sigma\varrho} A_{\lambda\zeta\eta 2} \\ &= \tfrac{1}{2}\varepsilon^{\alpha\beta\eta\zeta} A_{\alpha\beta\sigma\varrho} \left(\varepsilon^{\sigma\varrho 31} A_{3\zeta\eta 2} + \varepsilon^{\sigma\varrho 41} A_{4\zeta\eta 2}\right) \\ &= \varepsilon^{\alpha\beta\eta\zeta}\left(A_{\alpha\beta[24]} A_{3\zeta\eta 2} + A_{\alpha\beta[32]} A_{4\zeta\eta 2}\right) \\ &= \tfrac{1}{2}\varepsilon^{\alpha\beta\eta\zeta}\left(A_{\beta 24\alpha} A_{3\zeta\eta 2} + A_{\beta 32\alpha} A_{4\zeta\eta 2}\right) \\ &= \tfrac{1}{2}\varepsilon^{\alpha\beta\eta\zeta}\left(A_{2\beta\alpha 4} A_{3\zeta\eta 2} + A_{3\beta\alpha 2} A_{4\zeta\eta 2}\right) = 0 \end{aligned}$$

Hence $\quad t^{\mu}_{\nu} = 0 \quad , (\mu \neq \nu)$. (2.4)

---

2. For the meaning of the brackets enclosing indices, and any other symbols not here explicitly defined, see H.D.

Next, let $\mu = \nu = 1$, say, without loss of generality. Then

$$t^1_1 = \varepsilon^{\alpha\beta\eta\zeta}\varepsilon^{\sigma\rho\lambda 1}A_{\alpha\beta\sigma\rho}A_{\eta\zeta\lambda 1}$$

$$= 2\varepsilon^{\alpha\beta\eta\zeta}(\varepsilon^{3421}A_{\alpha\beta[34]}A_{\eta\zeta 21} + \varepsilon^{2431}A_{\alpha\beta[24]}A_{\eta\zeta 31} + \varepsilon^{2341}A_{\alpha\beta[23]}A_{\eta\zeta 41})$$

$$= \varepsilon^{\alpha\beta\eta\zeta}(A_{\alpha 34\beta}A_{\eta\zeta 21} + A_{\alpha 42\beta}A_{\eta\zeta 31} + A_{\alpha 23\beta}A_{\eta\zeta 41})$$

$$= \tfrac{1}{2}\varepsilon^{\alpha\beta\eta\zeta}(A_{\alpha 34\beta}A_{\eta 12\zeta} + A_{\alpha 24\beta}A_{\eta 13\zeta} + A_{\alpha 23\beta}A_{\eta 14\zeta})$$

$= \frac{1}{4}t$, as follows from the symmetry of the last expression,

remembering that $t = t^\alpha_\alpha$. Hence, in view of (2.4),

$$t^\mu_\nu = \tfrac{1}{4}\delta^\mu_\nu t \text{ , or } T^\mu_\nu \equiv 0 \text{ ,} \tag{2.5}$$

which was to be proved.

§3. — (a) Now in Weyl's geometry the conformal curvature tensor [Weyl(1)] $C_{\mu\nu\sigma\rho}$ has just the required symmetry properties. It is defined by

$$\overline{C}_{\mu\nu\sigma\rho} = \overline{B}_{\mu\nu\sigma\rho} + 2g_{[\nu[\rho}(\overline{G}_{\mu]\sigma]} - \tfrac{1}{6}g_{\mu]\sigma]}\overline{G}) \text{ ,} \tag{3.1}$$

where $\overline{B}_{\mu\nu\sigma\rho} = B_{\mu\nu\sigma\rho} - g_{\mu\rho}F_{\nu\sigma}$, (3.2)

and $\overline{G}_{\mu\nu} = \overline{B}_{\mu\nu\sigma}{}^{\sigma}$, $\overline{G} = \overline{G}_\mu{}^\mu$. (3.3)

$\overline{B}_{\mu\nu\sigma}{}^{\varepsilon}$ is Weyl's "direction-curvature", [Weyl(2)]. Notice

that <u>if</u> α,β <u>stands for any pair of the symbols</u> μ,ν,σ,ρ <u>then</u>

$$g^{\alpha\beta}\overline{C}_{\mu\nu\sigma\rho} = 0 \text{ .} \tag{3.4}$$

(b) Consider the tensor-density

$$\underline{t}^\mu_\nu = \varepsilon^{\alpha\beta\eta\zeta}\varepsilon^{\sigma\rho\lambda\mu}\overline{C}_{\alpha\beta\sigma\rho}\overline{C}_{\eta\zeta\lambda\nu} \tag{3.5}$$

of coordinate-weight +2, where $\varepsilon^{\mu\nu\sigma\rho}$ is the contravariant tensor-density of Levi-Civita, [Veblen(2)] . This is of the form (2.2). Raising the subscripts α,β,η,ζ we obtain [Veblen(3)]

$$t^\mu_\nu = g^{-1}\underline{t}^\mu_\nu = \delta^{\sigma\rho\lambda\mu}_{\alpha\beta\eta\zeta}\overline{C}^{\alpha\beta}{}_{\sigma\rho}\overline{C}^{\eta\zeta}{}_{\lambda\nu} \text{ .} \tag{3.51}$$

By (3.4) it is obvious that the only surviving terms on the right hand side of (3.51) are $4\overline{C}^{[\lambda\mu]}{}_{\sigma\rho}\overline{C}^{[\sigma\rho]}{}_{\lambda\nu}$,

whence
$$t^\mu_\nu = \overline{C}^{\lambda\sigma\rho\mu}\overline{C}_{\lambda\sigma\rho\nu} \text{ ,} \tag{3.6}$$

in view of the symmetry properties of $\overline{C}_{\mu\nu\sigma\rho}$. Substituting in this from (3.1), we get after some tedious rearranging

$$t^{\mu\nu} = L^{\mu\nu} - \tfrac{1}{6} g^{\mu\nu}(K_1 - 3K_2 + 3K_4) \quad , \tag{3.61}$$

where $L^{\mu\nu}$ is given by H.D., equ. (8.1), <u>viz.</u>

$$L^{\mu\nu} = GG^{(\mu\nu)} - G^{\cdot\mu}_{\cdot\varepsilon} G^{(\nu\varepsilon)} - G^{\cdot\nu}_{\cdot\varepsilon} G^{(\mu\varepsilon)} - 2B^{(\mu\nu)\sigma\rho} G_{(\sigma\rho)} + B_{\sigma\rho\alpha}{}^{\mu} B^{\sigma\rho\alpha\nu} - g^{\mu\nu} F_{\sigma\rho} F^{\sigma\rho} . \tag{3.62}$$

Therefore $\quad t = t^{\nu}_{\nu} = \tfrac{1}{3} K_1 - 2K_2 + K_3 + 10K_4 \quad ,$ (3.63)

whence $\quad T^{\mu\nu} = t^{\mu\nu} - \tfrac{1}{4} g^{\mu\nu} t = L^{\mu\nu} - \tfrac{1}{4} g^{\mu\nu} L \quad ,$ (3.7)

with $\quad L = L^{\nu}_{\nu} = K_1 - 4K_2 + K_3 + 12K_4 \quad .$ (3.71)

But, by H.D. §8., $T^{\mu\nu}$ is just the Hamiltonian derivative with respect to $g_{\mu\nu}$ of $-\tfrac{1}{2}L$, which is now seen to vanish identically, in virtue of (2.5). By H.D. equ. (8.4)., the Hamiltonian derivative with respect to k_μ also vanishes identically; and hence the result stated in paragraph 1 is proved.

<u>References.</u>

Buchdahl, H.A. (1) This thesis, page 267.

Lanczos, C. (1) Annals of Math. II., Vol. 39, No. 4, <u>p.</u> 342. 1938.

Veblen, O. (1) Invariants of Quadratic Differential Forms. Cambridge Tracts, No. 24, 1927. <u>p.</u> 4, <u>ff.</u>

(2) <u>l.c.</u> <u>p.</u> 25. (3) <u>p.</u> 5, (3.41) <u>and</u> <u>p.</u> 9, (8.4)

Weyl, H. (1) Mathem. Zeitschr. Vol. 2. (1918), <u>p.</u> 404.

(2) <u>l.c.</u>, <u>p.</u> 403.

ON A SET OF CONFORM-INVARIANT EQUATIONS OF THE GRAVITATIONAL FIELD.

by H.A. Buchdahl, Department of Physics, University of Tasmania.

§1. Eddington [1] has considered equations of the gravitational field in empty space which are of the fourth differential order, viz. the sets of equations which express the vanishing of the Hamiltonian derivatives of certain fundamental invariants. The author has shown [2] that a wide class of such equations are satisfield by any solution of the equations

$$G_{\mu\nu} = \lambda g_{\mu\nu} \quad , \tag{1.1}$$

where $G_{\mu\nu}$ and $g_{\mu\nu}$ are the components of the Ricci tensor and the metrical tensor repectively, whilst λ is an arbitrary constant. For a V_4 this applies in particular when the invariant referred to above is chosen from the set

$$\left.\begin{aligned} K_1 &= G^2 \; (= G^\mu_\mu G^\nu_\nu) \\ K_2 &= G_{\mu\nu} G^{\mu\nu} \\ K_3 &= B_{\mu\nu\sigma\varrho} B^{\mu\nu\sigma\varrho} \quad , \end{aligned}\right\} \tag{1.2}$$

where $B_{\mu\nu\sigma\varrho}$ is the covariant curvature tensor. K_3 has been included since, according to a result due to Lanczos [3] , its Hamiltonian derivative $P_3^{\mu\nu}$ is a linear combination of $P_1^{\mu\nu}$ and $P_2^{\mu\nu}$, i.e. of the Hamiltonian derivatives of K_1 and K_2 . In fact

$$P_3^{\mu\nu} = 4P_2^{\mu\nu} - P_1^{\mu\nu} \quad . \tag{1.3}$$

It appears therefore that the most general invariant which will give rise to quasi-linear fourth order equations may be taken to be

$$K_1 + aK_2 \quad , \tag{1.4}$$

where a is a constant.

The question of the general solutions of such equations is as yet unsolved, even in the case of static spherically symmetric fields, which despite its relative simplicity presents the greatest difficulties. In the present paper we shall be concerned with a special case of (1.4), viz. with the invariant

$$K = 3K_2 - K_1 , \qquad (1.5)$$

and we shall show that, if $P^{\mu\nu}$ is the Hamiltonian derivative of K, then the equations

$$P^{\mu\nu} = 0 \qquad (1.6)$$

possess as solutions any line-element representing a space conformal to an Einstein space.

Furthermore we shall show that the general solution of (1.6) in the case of the static spherically symmetric field may be written

$$ds^2 = \psi(r)\left[-\gamma^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\phi^2 + \gamma dt^2\right] , \qquad (1.7)$$

where $\gamma = 1 - 2m/r - \lambda r^2/3$, (m,λ arbitrary constants), and $\psi(r)$ is an arbitrary function of r. We are of course dealing with a V_4 throughout.

§2. Let $C_{\mu\nu\sigma\varrho}$ be the conformal curvature tensor [4]

$$C_{\mu\nu\sigma\varrho} = B_{\mu\nu\sigma\varrho} - 2g_{[\mu[\nu}(G_{\sigma]\varrho]} - \tfrac{1}{6}g_{\sigma]\varrho]}G) . \qquad (2.1)$$

(For the meaning of brackets enclosing indices, vide Schouten [5])

Consider the invariant

$$\bar{K} = C_{\mu\nu\sigma\varrho}C^{\mu\nu\sigma\varrho} . \qquad (2.2)$$

Using (2.1) it is not difficult to show that we can write

$$\bar{K} = \tfrac{1}{3}K_1 - 2K_2 + K_3 . \qquad (2.3)$$

Now let

$$L = K_1 - 4K_2 + K_3 . \qquad (2.4)$$

Inserting this in (2.3) the latter becomes

$$\bar{K} = L - \tfrac{2}{3}K_1 + 2K_2 . \qquad (2.5)$$

In virtue of (1.3) the Hamiltonian derivative of L vanishes identically. Accordingly we consider simply the invariant K as given by (1.5). The Hamiltonian derivative $P^{\mu\nu}$ of K will be the same as that of $\overline{K}$, except for a trivial numerical factor.

§3. Consider the integral

$$J = \int K\sqrt{-g}\, d\tau \quad , \tag{3.1}$$

where K and g are formed with repect to a metrical tensor $g_{\mu\nu}$. In a conformal transformation in which the $g_{\mu\nu}$ are replaced by $\sigma g_{\mu\nu}$, where σ is an arbitary function of the coordinates, K becomes multiplied by σ^{-2} and $\sqrt{-g}$ by σ^2, (in a V_4!). Now $P^{\mu\nu}$ is defined by the equation

$$\delta J = \int P^{\mu\nu}\delta g_{\mu\nu}\sqrt{-g}\, d\tau \quad , \tag{3.2}$$

where variations vanish on the boundary of the region of integration. It follows that in a conformal transformation the $P^{\mu\nu}$ merely become multiplied by σ^{-3}. The set of equations

$$P^{\mu\nu} = 0 \tag{3.3}$$

therefore transforms into itself, and we accordingly say that it is <u>conform-invariant</u>. (Strictly speaking $P^{\mu\nu}$ itself is " conform- covariant"). Obviously therefore, if $\overline{g}_{\mu\nu}$ is a particular solution of (3.3), then the product of $\overline{g}_{\mu\nu}$ with an arbitrary function of the coordinates is a more general solution. Making use of the known results stated in §1 we therefore have that if $\overline{g}_{\mu\nu}$ is the metrical tensor of an Einstein space, [i.e. $\overline{g}_{\mu\nu}$ is the solution of the equations $G_{\mu\nu} = \lambda g_{\mu\nu}$, with arbitrary (constant) λ], and $\Lambda(x_1,x_2,x_3,x_4)$ is an arbitrary function of the coordinates, then

$$g^{*}_{\mu\nu} = \Lambda(x_i)\bar{g}_{\mu\nu} \qquad (3.5)$$

satisfies the set of equations (3.3), which proves our first assertion.

§ 4. Although it is not essential for our purpose it may be of interest to write down the explicit form of $P^{\mu\nu}$. In fact, using some results due to the author [6] , we find without difficulty, that

$$P^{\mu\nu} = S^{\mu\nu} - \tfrac{1}{4}g^{\mu\nu}S \quad , \qquad (4.1)$$

where

$$S^{\mu\nu} = 3\Box G^{\mu\nu} - G^{;\mu\nu} + 2GG^{\mu\nu} - 6B^{\mu\nu\sigma\rho}G_{\sigma\rho} \quad . \qquad (4.2)$$

(4.1) may also be written

$$\tfrac{1}{3}P^{\mu\nu} = 2C^{\mu\nu\sigma\rho}{}_{;\sigma\rho} - C^{\mu\nu\sigma\rho}G_{\sigma\rho} \quad . \qquad (4.3)$$

By (4.1) the spur P of $P^{\mu\nu}$ <u>vanishes identically</u>. [This is obviously necessary, since the spur of the Hamiltonian derivative of <u>any</u> conform-invariant action density $K\sqrt{-g}$ must vanish. This is easily proved by considering the special variation

$$\delta g_{\mu\nu} = \varepsilon g_{\mu\nu} \qquad (4.4)$$

in the equation of definition of Hamiltonian derivatives (<u>cf.</u> equ. (3.2)), where ε is an arbitrary infinitesimal function of the coordinates, vanishing on the boundary of the region of integration].

§5. We now come to the case of static spherically symmetric solutions of (3.3). Disregarding trivial arbitrary constants, the only Einstein spaces having the required property are [7]

$$\left.\begin{aligned} ds^2 &= \bar{g}_{\mu\nu}dx^{\mu}dx^{\nu} \equiv -\gamma^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta\, d\phi^2) + \gamma dt^2 \,, \\ \text{where} \quad \gamma &= 1 - 2m/r - \lambda r^2/3 \quad , \end{aligned}\right\} \qquad (5.1)$$

m and λ being constants of integration. (We consider 'different' solutions obtainable from one another by coordinate transformations as constituting the <u>same</u> solution).

By a suitable choice of coordinates every spherically symmetric static line element ~~maty~~ may be brought into the form

$$\left.\begin{array}{c} ds^2 = -e^{\mu}dr^2 - r^2(d\theta^2 + \sin^2\theta \, d\phi^2) + e^{\nu}dt^2 \;, \\ (\mu = \mu(r) \;,\; \nu = \nu(r) \quad). \end{array}\right\} \quad (5.2)$$

In the present case we may further simplify (5.2) by carrying out first a conformal transformation in which ds^2 is multiplied by $e^{-\nu}$ throughout, followed by a coordinate transformation $r \longrightarrow r'$, where r' is such that $r^2 \exp(-\nu(r)) = r'^2$. Omitting primes we need then only consider line-elements of the form

$$ds^2 = - e^{\mu(r)}dr^2 - r^2(d\theta^2 + \sin^2\theta \, d\phi^2) + dt^2 \,. \qquad (5.3)$$

It is not difficult to confirm that the components of the curvature tensor do not contain the <u>second</u> derivative of $\mu(r)$, and consequently $P^{\mu\nu}$ does not involve the <u>foubth</u> derivative of $\mu(r)$.

Consider now the identity

$$P^{\mu\nu}{}_{;\nu} = 0 \;, \qquad (5.4)$$

which is satisfied because $P^{\mu\nu}$ is a Hamiltonian derivative [8]. Since $G^{\sigma\tau} \equiv 0$, $(\sigma \neq \tau)$, we also have $P^{\sigma\tau} \equiv 0$, $(\sigma \neq \tau)$, so that on chosing $\sigma = 1$ in (5.4), (the other three possiblities are trivial), we have $P^{1\tau}{}_{;\tau} = 0$, or

$$\frac{\partial P^{11}}{\partial r} \equiv - \sum_{\varepsilon} \{\varepsilon\varepsilon, 1\} P^{\varepsilon\varepsilon} - \sum_{\varepsilon} \{1\varepsilon, \varepsilon\} P^{11} \quad . \qquad (5.5)$$

Now the right hand side of this equation does not involve the fourth derivative of $\mu(r)$, so that P^{11} cannot contain the third derivative of $\mu(r)$. Hence the first equation of (3.3), <u>viz</u>.

$$P^{11} = 0 \;, \qquad (5.6)$$

is an ordinary non-linear differential equation of the <u>second</u> order for $\mu(r)$, the general solution of which involves two arbitrary constants. This solution we however already know: it is the function in (5.3) which arises from (5.1) when we subject the latter to the two transformations described above. It follows at once that all static spherically symmetric solutions of (3.3) can be written in the form (1.7); which was to be proved.

§6. (a) It would be of interest to examine whether the results above could be applied to the question of the definitions of clocks and measuring rods in general relativity theory. For, clearly, considering for the sake of simplicity the empty space surrounding the sun, the gravitational field there can always be represented in the form (1.7), no matter how strange a "clock" we may happen to be using - 'strange' within certain limits, at any rate. Once $\psi(r)$ has been fixed, (its form being determined by the behaviour of the particular clock employed), the field equations have a unique solution, according to the result of the previous section.

(b) It may be noted that just the set of equations (4.1), (4.2) is obtained in Weyl's theory [9] if we attempt to set up field equations by choosing

$$\delta\int \bar{C}_{\mu\nu\sigma\varrho}\,\bar{C}^{\mu\nu\sigma\varrho}\sqrt{-g}\,d\tau = 0 \qquad (6.1)$$

as the determining gauge-invariant action principle. This is of course not surprising, since Weyl's conformal curvature tensor $\bar{C}_{\mu\nu\sigma\varrho}$ [10] does not involve the 'electro-magnetic potentials' at all. But it is interesting to observe that in this case we can at least obtain convergent solutions of the field equations, (cf. [11]). On the other hand the 'unity' of gravitation and electricity is then of an even more dubious kind.

References.

1. Eddington, A.S., The Mathematical Theory of Relativity, 2nd. edn. Cambridge University Press, 1930. §62, p. 141.

2. Buchdahl, H.A., Proc.Math.Soc. Edinburgh, In the press. and

   Proc. Nat. Acad., Vol.38, No.2, p.66. 1948. (This thesis, page 236 and 252 respectively)

3. Lanczos, C., Annals of Math.,II, Vol.39, No.4, p.842. 1938.

4. Eisenhart, L.P., Riemannian Geometry. Princeton University Press 1926. Ch.II., p. 90.

5. Schouten, J.A., Der Ricci-Kalkül. Springer, Berlin. 1924. pp. 25 and 31.

6. Buchdahl, H.A., Oxford Quarterly Journal of Mathematics, In the press. (This thesis, p.242).

7. Eddington, A.S. as [1], §45, p. 100.

8. Eddington, A.S. as [1], §61, p. 140.

9. Weyl, H., Space, Time, Matter. Methuen, 1922. §§35-36.

10. Weyl, H., Math. Ztschr. Vol.2, 1918, p. 404.

11. Bergmann, P., Introduction to the Theory of Relativity. Prentice Hall. New York. 1946. Ch.XVI., p.253.

ON THE PRINCIPLE OF CARATHÉODORY.

by H. A. Buchdahl,

Physics Department, University of Tasmania.

§1. In most treatments of thermodynamics the Second Law is stated in one or other of the original forms due to Clausius and Thomson. Carathéodory, however, in his axiomatic development of thermodynamics, replaced the traditional statements of the Second Law by what has become known as the Principle of Carathéodory. A more widespread knowledge of the methods of Carathéodory[1.] seems desirable, not least because of their great didactic value. Experience shows that they can be understood by undergraduates in the second or third year of a course in physics. These methods involve, however, one mathematical theorem (the Theorem of Carathéodory) the usual proofs[2.] of which are often so unpalatable to the physicist, that the theorem itself may form a serious obstacle to a proper understanding of the whole treatment. We shall therefore give in a subsequent

---

1. Carathéodory, C., Math. Annalen. Vol.67, p.355, 1909 - and Sitz.d.Preu. Akad.d.Wiss. p.39. 1925.
2. Born, M., Phys. Ztschr. XXII. p.251. - and Chandrasekhar, S., Stellar Structure, Ch.I , p.21. 1939 (Uni. Chicago Press) - and Carathéodory, C. loc.cit., p.369. (1909).

paper an alternative proof of this theorem (for the case of three variables) which may be found more attractive. It is desirable in the development of any physical theory that there should be a clear-cut division between empirical content and mathematical method. In the case we are considering, once Carathéodory's Theorem is understood, as a theorem in pure mathematics, the existence of a certain single-valued function of the variables of state is at once seen to be an <u>immediate</u> consequence of the (generalised) empirical knowledge which is contained in the Second Law (in the form of Carathéodory's Principle). In the usual treatments the necessity for the existence of this function is often hidden behind a welter of abstract engines and cycles which seems to leave most students without much appreciation of what has been proved, and without any clear understanding of the phenomenological meaning of entropy.

§2. Keeping the preceding remarks in mind, and in order to emphasise the close analogy which exists between the 'physical argument' and the 'mathematical argument' of this treatment we propose to deal in this paper with some general considerations concerning the Principle of Carathéodory. The latter may be stated as follows:

<u>In the neighbourhood of any arbitrary initial state J_0 of a physical system there exist neighbouring states J which are not accessible from J_0 along adiabatic paths.</u>

This principle thus takes as a starting point the (empirical) recognition that if two states, $\underline{J_o}$ and $\underline{J}$, of a given adiabatically enclosed[3] (thermodynamic) system be prescribed, and granted (i) the purely mechanical possibility of the transition from $\underline{J_o}$ to $\underline{J}$, and (ii) that such a transition would not violate the demands which the First Law of Thermodynamics already imposes upon it: then the transition from $\underline{J_o}$ to $\underline{J}$ may nevertheless be impossible, whilst at the same time the reverse transition is possible. We then say that the thermodynamic weight of $\underline{J_o}$ exceeds that of $\underline{J}$.

Let us consider an elementary example. If $\underline{J_o} \rightarrow \underline{J}$ stands for the phrase 'the transition of the system from the state $\underline{J_o}$ to the state $\underline{J}$ ', let $\underline{J_o(h_o,t_o)}$ be the state of the Joule paddle-wheel apparatus, the contents of the calorimeter being at temperature $\underline{t_o}$, and the mass $\underline{m}$ at height $\underline{h_o}$. Let $\underline{J(h,t)}$ be a second state of the system, where $\underline{t} < \underline{t_o}$ and $\underline{h} > \underline{h_o}$, in such a way that the energy difference of the contents of the calorimeter, to which corresponds the temperature difference $\underline{t_o} - \underline{t}$, is just accounted for by the potential energy difference of the mass ($= \underline{mg(h - h_o)}$) in the two states of the system respectively. Then we know empirically that $\underline{J_o}(\underline{h_o},\underline{t_o}) \rightarrow \underline{J}(\underline{h},\underline{t})$ is impossible, notwithstanding the fact that neither the First Law, nor the laws of mechanics would be violated in this transition; $J(h,t) \rightarrow \underline{J_o}(\underline{h_o},\underline{t_o})$

---

3. A system is said to be adiabatically enclosed if a state of equilibrium of the system can be disturbed only by mechanical means (i.e. such as stirring, shaking, or the passage of electric currents.)

is, however, possible.

§3. In order to simplify the following considerations, which are of a fairly general kind, let a system _K_ consist of a gas within an envelope[4], the volume occupied by the gas being _v_, at a pressure _p_. We regard the mechanical variables _p_ and _v_ as the independent _variables of state_, i.e. the quantities _p_ and _v_ _define_ the state of the system, and are (within certain limits) variable at will. In a manner which we need not consider here, the conditions for thermal equilibrium lead us to associate with given values of _p_ and _v_ a number _t_, such that two such systems _K_ and _K'_ can be in thermal equilibrium if, and only if, the corresponding numbers _t_ and _t'_ are equal[5]. That is, empirical knowledge concerning the thermal equilibrium of physical systems leads to the definition of a single-valued function _t(p,v)_ of the variables of state, which expresses a new property of the system, viz. the property of being or of not being in thermal equilibrium with another system when the two are brought into non-adiabatic contact. Any other such definitive property[6] (often expressed in

---

4. The envelope itself is not to be regarded as forming a part of the system.

5. The apparent existence of more than one value of _t_ for given _p_ and _v_ (e.g. water near 4°C.) would show the incorrectness of the assumption that _p,v_ constituted a _sufficient_ set of independent variables of state.

6. A property of this type evidently cannot depend on the previous history of the system.

the form of a natural law) may similarly be expected to lead to the attachment of a certain number to every given state, i.e. to the definition of a new single-valued function of the variables of state associated with the system, which expresses this property. The First Law of Thermodynamics is an excellent example; it generalises the result of a very great number of experiments in the statement that the mechanical work W done by a system in any adiabatic transition between two given states depends upon these states alone, i.e. not on the manner of transition. The definition of a new single-valued function of the variables of state, the energy U of the system, is an immediate consequence of this statement. (The term 'quantity of heat' (Q) then appears merely as an abbreviation for the difference between the actual work done in a given transition and the change in the value of the energy function which occurs in it. Thus if U_0, U are the values of the energy in the initial and final states respectively, then

$$Q = (U - U_0) + W \quad . \qquad (1))$$

§4. After the preliminary observations of the preceding section we return to the consideration of Carathéodory's Principle. As we have seen, the latter expresses a definitive property of the system, viz. the property that when adiabatically isolated, the possibility or impossibility of $J_0 \rightarrow J$ depends upon J_0 and J alone (subject to certain other well-defined conditions being already satisfied). Accordingly we may expect the principle to lead to a new single-valued function of the variables of state $\underline{S}$[7], suwh that $\underline{S}$ is a measure of the thermodynamic weight of the state $\underline{J}$. We call $\underline{S}$ the entropy of the system. It follows at once (the sign of $\underline{S}$ being suitably chosen) that $\underline{J}_0 \rightarrow \underline{J}$ is possible if $\underline{S} \geqslant \underline{S}_0$, and impossible if $\underline{S} \nless \underline{S}_0$; for the condition of accessibility cannot be expressed in any essentially different way in terms of a pair of numbers which must enter into the relations quasi-symmetrically. Moreover, let $\underline{S} > \underline{S}_0$; then $\underline{J}_0 \rightarrow \underline{J}$ is possible. But having effected $\underline{J}_0 \rightarrow \underline{J}$, $\underline{J} \rightarrow \underline{J}_0$ is now impossible, for now $\underline{S}(\text{final}) < \underline{S}(\text{initial})$. That is, $\underline{J}_0 \rightarrow \underline{J}$ is irreversible. Clearly $\underline{J}_0 \rightarrow \underline{J}$ is reversible only if $\underline{S}_0 = \underline{S}$. The last results may be summed up as follows:

A transition of an adiabatically enclosed system is impossible, possible reversibly, or possible irreversibly according as the entropy of the initial state is greater than, equal to, or less than that of the final state.

This at once gives rise to the corollary that <u>in any adiabatic transition of a system the entropy can never decrease</u>. This is the

---

7. It is conceivable that it could define more than one new function determining the mutual accessibility of different states. But it is difficult to see how this could come about on the basis of a law of the type under consideration. However, it appears that we must ultimately rely upon the confirmation obtained from a mathematical treatment of the problem.

so-called Principle of Increase of Entropy, which shows that unlike mass, energy, charge etc. entropy obeys a one-sided conservation law.

We shall not pursue the physical consequences of the Principle of Carathéodory beyond this point; for the elucidation of the phenomenological meaning of entropy as 'transition potential' has been dealt with at sufficient length for our purpose.

§5.(a) Finally we briefly examine how the considerations above indicate to us how to begin with the mathematical formulation of the consequences of Carathéodory's Principle. To do this it is sufficient to consider a system **L** with three independent variables of state [such as the systems **K** and **K'** of §3 in thermal equilibrium], which we take to be v, v' and the common temperature t. Now Carathéodory's Principle speaks of arbitrary adiabatic transitions. It applies therefore *a forteriori* to quasi-static[8.]

---

8. A transition of a system **L** is said to be *quasi-static* if in the course of it **L** passes through a continuous series of states of equilibrium. (This is equivalent to a reversible transition, which necessarily proceeds at an infinitesimal rate).

adiabatic transitions. During an infinitesimal part of it the work done by L is $pdv + p'dv'$; and since the transition is adiabatic this work must, in virtue of the definition of energy, be equal to the change dU in the energy $U(v,v',t)$ of L , i.e.

$$(\partial U/\partial v + p)dv + (\partial U/\partial v' + p')dv' + \partial U/\partial t . dt = 0 . \qquad (2)$$

Thus the quasi-static adiabatic transitions of L are subject to a condition of the form

$$(dQ \equiv) \quad P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz = 0 , \qquad (3)$$

where P,Q,R are certain functions[9.] of the independent variables x,y,z. Interpreting the latter as right angled cartesian coordinates of a 'picture space' A, every state of equilibrium of L is represented by a point in A. By definition a quasi-static transition must therefore be represented by a continuous curve C in A. If the transition is also adiabatic, C is restricted by the equation (3). In other words: The 'quasi-static adiabatics' of L are the solution curves of the differential equation (3).

---

9. Equation (3) preserves its form under any substitution of independent variables. Thus, if the set x,y,z be given as functions of the new set x',y',z', then the equation becomes $P'dx' + Q'dy' + R'dz' = 0$, where $P'(x',y',z') = P\partial x/\partial x' + Q\partial y/\partial x' + R\partial z/\partial x'$, etc. (Note that the Q here has of course nothing to do with the symbol for quantity of heat.)

 303

But in §4 we tentatively accepted the existence of a certain function <u>S</u>, and we saw that it remains constant in a quasi-static adiabatic transition. That is, <u>as a consequence of the Second Law there exists a function S such that the equation</u>

$$Pdx + Qdy + Rdz = 0 \tag{4}$$

<u>implies</u>

$$dS = 0 \quad , \tag{5}$$

<u>and which has the properties described in §4.</u> Hence if we are on the right track we may expect that there exists another function $\omega(x,y,z)$, such that[10.]

$$Pdx + Qdy + Rdz \equiv \omega dS \quad . \tag{6}$$

(b) We arrived at the tentative equations (5) and (6) by means of a direct physical 'plausibility argument' based on Carathéodory's Principle. These equations may now be put on a rigourous basis through an application of the Theorem of Carathéodory, the statement, and a new proof of which will form the substance of another paper[11.], in accordance with our original intention of delimiting the <u>mathematical</u> core of the consequences of the Second Law.

---

10. For quite unrestricted <u>P,Q,R</u> this is, in general, impossible; in fact the 'condition of integrability'
$P(\partial Q/dz - \partial R/\partial y) + Q(\partial R/\partial x - \partial P/\partial z) + R(\partial P/\partial y - \partial Q/\partial x) = 0$
must be satisfied. See <u>Forsyth A.R.</u>, Differential Equations, 3rd edn. 1903. pp.282 - 284. (MacMillan)
11. This thesis, page 304. (Amer. Jour. Physics, In the press).

ON THE THEOREM OF CARATHÉODORY.

by H. A. Buchdahl,

Physics Department, University of Tasmania.

§1.(a) The Second Law of Thermodynamics in the form of the Principle of Carathéodory states that if we consider different states of a given physical system in the neighbourhood of any arbitrary state J_o there are states J which are not accessible from J_o along adiabatic paths. This principle was considered by the author in a previous paper[1.] from a physical standpoint. As announced there, we now propose to give a straightforward analytical proof of the Theorem of Carathéodory[2.]. This is most easily stated in terms of a picture space $\underline{A}$, with rectangular coordinates $\underline{x},\underline{y},\underline{z}$ (of which a certain finite region $\underline{D}$ is contemplated). The theorem then takes the form: In the neighbourhood of any arbitrary point G_o there are points G which are not accessible from G_o along solution curves of the equation

$$P(x,y,z)dx + Q(x,y,z)dy + R(x,y,z)dz = 0 \tag{1}$$

if, and only if, the equation is integrable.

1. This thesis, page 295. (Am. J. Phys., In the press). References to this paper will be distinguished by the letter $\underline{P}$ throughout.
2. We restrict the theorem to the case of an equation in three variables.

The equation is called integrable if it is equivalent to a single finite algebraic relation

$$F(x,y,z) = \text{const.}\ , \tag{2}$$

that is, if there exist functions $\lambda(x,y,z)$, $F(x,y,z)$ such that[3.]

$$Pdx + Qdy + Rdz \equiv \lambda dF \ . \tag{3}$$

(b) The theorem, whose close verbal resemblance to the principle should be noted, allows of an immediate rigourous mathematical formulation of the consequences of the Second Law. For (in terms of the picture space A) the quasi-static adiabatics of a system L with three independent variables of state are the solution curves of a differential equation of just the form (1), (cf. P §5(a)). And if we consider quasi-static transitions of L the stated principle asserts: in the neighbourhood of any arbitrary point G_o there are points G which are not accessible from G_o along solution curves of (1). The theorem then immediately provides us with equation (3) : and this has already the form P(7). We return to these considerations briefly in §4.

§2. It is not difficult to see that the solution of (1) is a set of curves; for the equation merely prescribes that at any point G

---

3. If P,Q,R are quite unrestricted such functions do not in general exist. In fact, they exist only if the 'condition of integrability', viz.
$P(\partial Q/\partial z - \partial R/\partial y) + Q(\partial R/\partial x - \partial P/\partial z) + R(\partial P/\partial y - \partial Q/\partial x) = 0$
is satisfied identically.

a certain line element with components dx, dy, dz shall be perpendicular to the (given) vector with components P, Q, R at G. Hence any curve which is such that the tangent to it at any point is perpendicular to the vector (P, Q, R) at that point is a solution curve of (1).[4]

To proceed with the proof of the theorem it is convenient to carry out a certain change of variables, for which purpose we determine a pair of functions $u(x,y,z)$, $\mu(x,y,z)$, such that

$$\left.\begin{aligned} P &= \frac{1}{\mu}\frac{\partial u}{\partial x} \\ Q &= \frac{1}{\mu}\frac{\partial u}{\partial y} \quad . \end{aligned}\right\} \tag{4}$$

(Such functions can always be found. Thus u is a solution of the differential equation $$Q\frac{\partial u}{\partial x} - P\frac{\partial u}{\partial y} = 0 \quad , \tag{5}$$ in which z is regarded as a constant).

Writing $$\mu R - \frac{\partial u}{\partial z} = Z \quad , \tag{6}$$

(1) takes the form $$du + Zdz = 0 \quad . \tag{7}$$

It is assumed that in the range D P, Q, R and μ are such that u and Z are single-valued, finite and continuous functions of x, y, z and possess finite and continuous first partial differential coefficients with respect to x, y, z. In place of x, y, z we now adopt u, y, z as independent variables, (physically: change of variables of state!); hence

$$Z = Z(u, y, z) \quad . \tag{8}$$

4. See also the excellent discussion of Forsyth, A.R., Differential Equations, 6th edn. pp. 318-324., 1943. (MacMillan).

Corresponding to this change of variables we now use a picture space $\underline{A}'$ with rectangular coordinates $\underline{u},\underline{y},\underline{z}$. In accordance with the assumptions just made there will be a reversible one-to-one correspondence between the points in the range $\underline{D}$ of $\underline{A}$ and the points in the corresponding range $\underline{D}'$ of $\underline{A}'$, and moreover it suffices to prove the theorem of §1 for the case of equation (7).

§3. (a) If $\underline{H_o(u_o,y_o,z_o)}$ is the point in $\underline{D}'$ which corresponds to an arbitrary point $\underline{G}_o$ in $\underline{D}$, consider how the passage along a solution curve of (7) from $\underline{H}_o$ to a neighbouring point $\underline{H}$ may actually be effected.

(i) First, pass in the plane $u = u_o$ from $\underline{H}_o$ to the point $\underline{H}_1$. Since, by (7), $\underline{z} = \underline{\text{const.}} = \underline{z_o}$ the coordinates of H_1 are

$$(u_o,y_1,z_o) \quad , \tag{9}$$

where y_1 may be chosen at will within $\underline{D}'$, ($-\sigma \leq y_1 - y_o \leq \sigma'$, say; $\sigma,\sigma' > 0$)

(ii) Next, pass in the plane $\underline{y} = \underline{y_1}$ from $\underline{H}_1$ to the point $\underline{H}_2$. (7) now reads

$$du + Z(u,y_1,z)dz = 0 \quad , \tag{10}$$

the solution of which may be written

$$u = q(z,y_1) \quad , \tag{11}$$

where the constant of integration is so chosen that

$$q(z_o,y_1) = u_o \quad . \tag{12}$$

Hence the coordinates of $\underline{H}_2$ are

$$(q(z_2,y_1),\ y_1,z_2) \quad , \tag{13}$$

where $\underline{z}_2$ may be chosen at will within $\underline{D}'$.

(iii) Finally, pass in the plane $\underline{z} = \underline{z}_2$ from $\underline{H}_2$ to the point $\underline{H}$, which will have the coordinates

$$(q(z_2,y_1),\ y_3,z_2) \quad , \tag{14}$$

where $\underline{y}_3$ may be chosen at will within D' .

(b) Now $q(z,y_1)$ is a finite and continuous function of $\underline{z, y_1}$; moreover, $\frac{\partial q}{\partial y_1}$ is finite and continuous in $\underline{D}'$.[5] Consider the equation

$$\frac{\partial q}{\partial y_1} = 0 \quad . \tag{15}$$

Two possibilities arise: either (15) is satisfied identically everywhere in $\underline{D}'$ or it is not. In the latter case, [equ.(7), and therefore (1), is then, of course, not integrable] , we choose $\underline{H}_o$ such that

$$\left(\frac{\partial q}{\partial y_1}\right)_{y_1 = y_o} \neq 0 \quad . \tag{16}$$

Then, - keeping the continuity conditions in mind, - there exist positive numbers ε_1 , ε_2 such that for any $\underline{z}_2$,$(z_o - \varepsilon_1 \leq z_2 \leq z_o + \varepsilon_1)$, we can determine y_1 , $(y_o - \sigma \leq y_1 \leq y_o + \sigma')$, so that $q(z_2,y_1)$ takes on any prescribed value lying between the limits $u_o - \varepsilon_2$, $u_o + \varepsilon_2$. Hence if, for some positive number ε_3 , we take

$$y_o - \varepsilon_3 \leq y_3 \leq y_o + \varepsilon_3 \ ,$$

we see at once that certainly all points $\underline{H}$ lying in the range

---

5. Kamke, E., Differentialgleichungen, 2nd edn. p.35., 1943. (Becker, Erler; Leipzig).

$$\left.\begin{array}{r} -\varepsilon_2 \leq u - u_0 \leq \varepsilon_2 \\ -\varepsilon_3 \leq y - y_0 \leq \varepsilon_3 \\ -\varepsilon_1 \leq z - z_0 \leq \varepsilon_1 \end{array}\right\} \qquad (17)$$

are accessible from $\underline{H}_0$ along solution curves of (7), (viz. in the manner specified in §3(a)). ε_1, ε_2, ε_3 must of course be such that the various points considered above lie in $\underline{D}'$.

(c) From the last result it follows immediately that under the conditions stated there in general inaccessible points in the neighbourhood of any arbitrarily chosen initial point only if (15) is satisfied identically in $\underline{D}'$. In that case $\underline{q}$ and therefore $\underline{Z}$ are independent of $\underline{y}_1$, (i.e. of $\underline{y}$), and (7) is obviously integrable. Consequently (1) is integrable; so that, by (2), all solution curves passing through $\underline{G}_0$ lie in the surface

$$F(x,y,z) = F(x_0,y_0,z_0) \quad , \qquad (18)$$

for at any point a small displacement must, as we have seen, take place perpendicularly to $\underline{P}, \underline{Q}, \underline{R}$, i.e. perpendicularly to the normal to the surface at the point. Accordingly all points $\underline{G}$ in the neighbourhood of $\underline{G}_0$ which do not lie on the integral surface (18) are inaccessible from $\underline{G}_0$ along solution curves of (1); and hence the theorem is proved.

§4. We have now established firmly that for quasi-static transitions of the system

$\underline{L}$ $$dQ = \lambda dF \quad , \qquad (19)$$

where λ and $\underline{F}$ are functions of the variables of state, [cf. equs. (3) and P(4)]. The paper of Born[6.] may be consulted for a simple demonstration that (19) may be written in the usual form

$$dQ = TdS \quad , \tag{20}$$

where $T(t)$ is a universal function[7.] of the thermometric temperature t of L:

$$\frac{d \log T(t)}{dt.} = \frac{\partial \log \lambda}{\partial t} \quad , \tag{21}$$

whilst

$$S = \int \Phi(F)\, dF \quad , \tag{22}$$

where $\Phi(F)$ is some function of F. In the same paper it is shown that S then indeed possesses all the properties which we expect it to have on the basis of the discussion of P§(4); and the question of the practical determination of the various functions introduced is briefly considered.

---

6. Born, M., Phys. Ztschr. XXII. pp. 282-286. 1931.
7. Universal in the sense that, if L' is another system in equilibrium with L, then $\partial \log \lambda'/\partial t = \partial \log \lambda/\partial t$, where λ' is the 'integrating denominator' of dQ', dashed and undashed quantities referring to L' and L respectively. (Note that t is the same for both systems).